

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 11, 2005, 10:08:01 ; Search time 217 Seconds  
(without alignments)  
3804.014 Million cell updates/sec

Title: US-10-783-417-1  
Perfect score: 2208  
Sequence: 1 atgacatacaataacgataa.....atcatcaaaaacacttga 2208

Scoring table: IDENTITY NUC  
Gapop 10.0, Gapext 1.0

Searched: 3392430 seqs, 186927314 residues

Total number of hits satisfying chosen parameters: 6784860

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

Published Applications NA New:  
1: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq.\*  
2: /cgn2\_6/ptodata/2/pubpna/US06\_NEW\_PUB.seq.\*  
3: /cgn2\_6/ptodata/2/pubpna/US07\_NEW\_PUB.seq.\*  
4: /cgn2\_6/ptodata/2/pubpna/US08\_NEW\_PUB.seq.\*  
5: /cgn2\_6/ptodata/2/pubpna/PC1\_NEW\_PUB.seq.\*  
6: /cgn2\_6/ptodata/2/pubpna/US10\_NEW\_PUB.seq.\*  
7: /cgn2\_6/ptodata/2/pubpna/US11\_NEW\_PUB.seq.\*  
8: /cgn2\_6/ptodata/2/pubpna/US11\_NEW\_PUB.seq3.\*  
9: /cgn2\_6/ptodata/2/pubpna/US11\_NEW\_PUB.seq3.\*  
10: /cgn2\_6/ptodata/2/pubpna/US60\_NEW\_PUB.seq.\*

Prod. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	69	3.1	173602	7	US-11-121-086-25
2	63.8	2.9	171486	7	US-11-121-086-105
3	60.8	2.8	2031	7	US-11-058-727-51
4	60.8	2.8	2031	7	US-11-058-727-83
5	60.8	2.8	2031	7	US-11-108-389-51
6	60.8	2.8	2031	7	US-11-108-389-83
7	60.6	2.7	173602	7	US-11-121-086-25
8	59.6	2.7	2028	7	US-11-058-727-35
9	59.6	2.7	2028	7	US-11-058-727-71
10	59.6	2.7	2028	7	US-11-108-389-39
11	59.6	2.7	2028	7	US-11-108-389-71
12	58.2	2.6	2022	7	US-11-058-727-13
13	58.2	2.6	2022	7	US-11-108-389-13
14	58.2	2.6	3633	7	US-11-058-727-3
15	58.2	2.6	3633	7	US-11-108-389-3
16	58.2	2.6	6613	7	US-11-058-727-18
17	58.2	2.6	6613	7	US-11-108-389-18
18	58.2	2.6	139054	7	US-11-121-086-96
19	58	2.6	2025	7	US-11-058-727-41
20	58	2.6	2025	7	US-11-058-727-73
21	58	2.6	2025	7	US-11-108-389-41
22	58	2.6	2025	7	US-11-108-389-73
23	57.6	2.6	2019	7	US-11-058-727-55

24	57.6	2.6	2019	7	US-11-058-727-59	Sequence 59, Appl
25	57.6	2.6	2019	7	US-11-058-727-87	Sequence 87, Appl
26	57.6	2.6	2019	7	US-11-058-727-91	Sequence 91, Appl
27	57.6	2.6	2019	7	US-11-108-389-55	Sequence 55, Appl
28	57.6	2.6	2019	7	US-11-108-389-59	Sequence 59, Appl
29	57.6	2.6	2019	7	US-11-108-389-87	Sequence 87, Appl
30	57.6	2.6	2019	7	US-11-108-389-91	Sequence 91, Appl
31	57	2.6	171486	7	US-11-121-086-105	Sequence 105, Appl
32	56.8	2.6	151169	7	US-11-121-086-38	Sequence 38, Appl
33	55.2	2.5	2022	7	US-11-058-727-49	Sequence 49, Appl
34	55.2	2.5	2022	7	US-11-058-727-81	Sequence 81, Appl
35	55.2	2.5	2022	7	US-11-108-389-49	Sequence 49, Appl
36	55.2	2.5	2022	7	US-11-108-389-81	Sequence 81, Appl
37	54.8	2.5	2025	7	US-11-058-727-47	Sequence 47, Appl
38	54.8	2.5	2025	7	US-11-058-727-79	Sequence 79, Appl
39	54.8	2.5	2025	7	US-11-108-389-47	Sequence 47, Appl
40	54.8	2.5	2025	7	US-11-108-389-79	Sequence 79, Appl
41	54.4	2.5	2019	7	US-11-058-727-57	Sequence 57, Appl
42	54.4	2.5	2019	7	US-11-058-727-89	Sequence 89, Appl
43	54.4	2.5	2019	7	US-11-108-389-57	Sequence 57, Appl
44	54.4	2.5	2019	7	US-11-108-389-89	Sequence 89, Appl
45	54.4	2.5	2022	7	US-11-058-727-21	Sequence 21, Appl

#### ALIGNMENTS

RESULT 1									
US-11-121-086-25/c									
; Sequence 25, Application US/11121086									
; Publication No. US20050266459A1									
; GENERAL INFORMATION:									
; APPLICANT: NIELSEN, KIRSTEN V.									
; TITLE OF INVENTION: NUCLEIC ACID PROBES AND NUCLEIC ACID ANALOG PROBES									
; FILE REFERENCE: 09138.6000-00000									
; CURRENT FILING DATE: 2005-05-04									
; PRIOR APPLICATION NUMBER: 60/567,570									
; PRIOR FILING DATE: 2004-05-04									
; NUMBER OF SEQ ID NOS: 107									
; SOFTWARE: PatentIn version 3.3									
; SEQ ID NO 25									
; LENGTH: 173602									
; TYPE: DNA									
; ORGANISM: Homo sapiens									
US-11-121-086-25									
Query Match 3.1%; Score 69; DB 7; Length 173602;									
Best Local Similarity 45.7%; Pred. No. 0.0013;									
Matches 430; Conservative 0; Mismatches 500; Indels 11; Gaps 5;									
QY	585	TTAAATACGATTGGAATGTCACAAATGATTATTGAGAAATACCTGTTCCAACT	644						
DB	136987	TAT	136928						
QY	645	TGAACCTATATAACGCTATACCTATTTAGCGGAGCTGATATTTCATTAA	704						
DB	136927	AATATATTAACATTTAT	136668						
QY	705	TTTATTTACAAAGGTCGATGGCTGATGAATGGAATGCAGATATACATCTTCA	764						
DB	136867	TAT	136808						
QY	765	AATGAACCTATATGCTGGAACATCAGATGACTATTTAAACTTTAAAGAAATATAC	824						
DB	136807	AATATATTAAT	136748						
QY	825	TAAATATATGTAATCTGTTGGAATATCTATAGAAACAGACTTAAATCTTGAAGACA	884						
DB	136747	TAT	136688						
QY	885	ACCAATATGAAATGAGCTATATTTAATGACTATGAGATATATGACCATTACTGAT	944						

[illegible]

```

RESULT 2
US-11-121-086-105
; Sequence 105, Application US/11/121086
; Publication No. US20050266459A1
;
GENERAL INFORMATION:
;
APPLICANT: NIELSEN, TIM S.
;
TITLE OF INVENTION: NUCLEIC ACID PROBES AND NUCLEIC ACID ANALOG PROBES
;
FILE REFERENCE: 09138.6000-00000
;
CURRENT APPLICATION NUMBER: US/11/121,086
;
CURRENT FILING DATE: 2005-05-04
;
PRIOR APPLICATION NUMBER: 60/567,570
;
PRIOR FILING DATE: 2004-05-04
;
NUMBER OF SEQ ID NOS: 107
;
SOFTWARE: PatentIn version 3.3
;
SEQ ID NO 105
;
LENGTH: 171486
;
TYPE: DNA
;
ORGANISM: Homo sapiens
;
US-11-121-086-105

```

Query Match	2.9%;	Score 63.8;	DB 7;	Length 171486;
Best Local Similarity	44.0%;	Pred. No. 0.011;		
Matches 410;	Conservative 0;	Mismatches 512;	Indels 9;	

QY 690 TAATTTTCATTAAATTTATTTCAACAAGTCTGTAATGGCTGATGAATGGAATGCAGA 749  
Db 16132 TAATATATATTTATTAATAATATATATCTTATATATATTAATATATATATATAGATATTAAT 16191

[illegible]

RESULT 3  
US-11-058-727-51  
; Sequence 51, Application US/11058727  
; Publication No. US20050261483A1

APPLICANT: Andre R. Adad  
APPLICANT: Ronald D. Flannagan  
APPLICANT: Rafael Herrmann  
APPLICANT: Theodore W. Kahn  
APPLICANT: Albert L. Lu  
APPLICANT: Billy Fred McCutchen  
APPLICANT: James K. Presnail

```

; APPLICANT: James F.H. Wong
; APPLICANT: Cao-Guo Yu
; TITLE OF INVENTION: Genes Encoding Proteins With Pesticidal
; TITLE OF INVENTION: Activity
; FILE REFERENCE: 35718/287809
; CURRENT APPLICATION NUMBER: US/11/058,727
; CURRENT FILING DATE: 2005-02-15
; PRIOR APPLICATION NUMBER: 60/391,786
; PRIOR FILING DATE: 2002-06-26
; PRIOR APPLICATION NUMBER: 60/460,787
; PRIOR FILING DATE: 2003-04-04
; PRIOR APPLICATION NUMBER: 10/606,320
; PRIOR FILING DATE: 2003-06-25
; NUMBER OF SEQ ID NOS: 134
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO: 51
; LENGTH: 2031
; TYPE: DNA
; ORGANISM: Bacillus thuringiensis (mutated)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(2031)
; US-11-058-727-51

```

```

Query Match      2.8%; Score 60.8; DB 7; Length 2031;
Best Local Similarity 47.1%; Pred. No. 0.014;
Matches 292; Conservative 0; Mismatches 292; Indels 36; Gaps 2;

```

```

QY 367 GACAAAACGATGACACAACTTTTAAATGGGAAATTTTGTGATACACCGTTA 426
DB 332 GAAAAGAGTCATGGGAAATTTTATGGAACAAAGTGAAGAACTCAATTAATTAATA 381
QY 427 ACAGAAAGATTAACAGCTAAAGTTACAACTTTAGAGATTTAGACAAATATTACAA 486
DB 382 GCAGAAATTCAGAGAAATTAAGCGCTTTCGAAATTAAGAGATTAAGTAAATTAACA 441
QY 487 AGCTAAATTAACAGATTAAGTATGAGAAATTTAAAGAACTCAAGCTCTGGATTA 546
DB 442 TTATATCTAACTGGCGCTTGAAGAAATGGGAAATTCATTTGAGTGAAGTATGTTTA 501
QY 547 CCACCATCATGACATTAACACAGCTGCTTGACTCTTAAATACGATTGAGAAATGTT 606
DB 502 AATGCTTCCGCGCAGCTTACGAGATG-----TGCAGAAATCGATTGAAATCCCTG 552
QY 607 CACATGATTTTATTCGAGAAATACCTGTTTCAACTGGAATTAATAAGCTATTA 666
DB 553 GATAGTTTATTTACGCAATTAATATGCGCATCTTTAGAGTACAAATTTGAACTACCATTC 612
QY 667 CTACCTATTTATGCGCAAGCTGCTAATTTTCAATTTTAAATTTTATTAACAAGTGTGAA 726
DB 613 CTTACTGATATGCAATGAGCAACCTTCACTTTTACTGTTTAAAGAGCGGTCAATT 672
QY 727 TTGCTGATGAATGGAATGCAATATACATCTTCAAAATTTGAACCTTAATCTGGAACA 786
DB 673 TTTGAGAGAAATGGGGAATGTCAACAAT-----ACT 705
QY 787 TCAGATGACTATTATAAATTTTAAAGAAATATTAATTAATAGTAACTATGTA 846
DB 706 ATTAATTAATTAATTAATGATGTCAAATGAAACTTACTGCAATTAATTTGATCACTGTGA 765
QY 847 AATACCTATAGAACAGACTAATAAATCTTAGAGCAACCAATATGAAATGAGATTA 906
DB 766 AAGTGATATGAAATCGTTTACCAAAATTTAAAGGACAGAGCGCTTAACAAATGGGTGAC 825
QY 907 TTTAATGACTATGAGATATATGACCAATTAATGATTAATGATCAATCTCTCAATTTTCT 966
DB 826 TATAACCAATTCCTGATAGAAATGACACGCGGTTTATGATGTTGCTGATTAATCCCA 885
QY 967 TTATATGATTAATAAAGATA 986
DB 886 AATTATGACACAGCAGCTA 905

```

```

RESULT 4
US-11-058-727-83
; Sequence 83, Application US/11058727
; Publication No. US20050261483A1
; GENERAL INFORMATION:
; APPLICANT: Andre R. Abad
; APPLICANT: Ronald D. Flanagan
; APPLICANT: Rafael Herrmann
; APPLICANT: Theodore W. Kahn
; APPLICANT: Albert L. Lu
; APPLICANT: Billy Fred McCutchen
; APPLICANT: James K. Presnall
; APPLICANT: James F.H. Wong
; APPLICANT: Cao-Guo Yu
; TITLE OF INVENTION: Genes Encoding Proteins With Pesticidal
; TITLE OF INVENTION: Activity
; FILE REFERENCE: 35718/287809
; CURRENT APPLICATION NUMBER: US/11/058,727
; CURRENT FILING DATE: 2005-02-15
; PRIOR APPLICATION NUMBER: 60/391,786
; PRIOR FILING DATE: 2002-06-26
; PRIOR APPLICATION NUMBER: 60/460,787
; PRIOR FILING DATE: 2003-04-04
; PRIOR APPLICATION NUMBER: 10/606,320
; PRIOR FILING DATE: 2003-06-25
; NUMBER OF SEQ ID NOS: 134
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO: 83
; LENGTH: 2031
; TYPE: DNA
; ORGANISM: Bacillus thuringiensis
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(2031)
; US-11-058-727-83

```

```

Query Match      2.8%; Score 60.8; DB 7; Length 2031;
Best Local Similarity 47.1%; Pred. No. 0.014;
Matches 292; Conservative 0; Mismatches 292; Indels 36; Gaps 2;

```

```

QY 367 GACAAAACGATGACACAACTTTTAAATGGGAAATTTTGTGATACACCGTTA 426
DB 332 GAAAAGAGTCATGGGAAATTTTATGGAACAAAGTGAAGAACTCAATTAATTAATA 381
QY 427 ACAGAAACATTAACAGCTAAAGTTACAACTTTAGAGATTTAGACAAATATTACAA 486
DB 382 GCAGAAATTCAGAGAAATTAAGCGCTTTCGAAATTAAGAGATTAAGTAAATTAACA 441
QY 487 AGCTAAATTAACAGATTAAGTATGAGAAATTTAAAGAACTCAAGCTCTGGATTA 546
DB 442 TTATATCTAACTGGCGCTTGAAGAAATGGGAAATTCATTTGAGTGAAGTATGTTTA 501
QY 547 CCACCATCATGACATTAACACAGCTGCTTGACTCTTAAATACGATTGAGAAATGTT 606
DB 502 AATGCTTCCGCGCAGCTTACGAGATG-----TGCAGAAATCGATTGAAATCCCTG 552
QY 607 CACATGATTTTATTCGAGAAATACCTGTTTCAACTGGAATTAATAAGCTATTA 666
DB 553 GATAGTTTATTTACGCAATTAATATGCGCATCTTTAGAGTACAAATTTGAACTACCATTC 612
QY 667 CTACCTATTTATGCGCAAGCTGCTAATTTTCAATTTTAAATTTTATTAACAAGTGTGAA 726
DB 613 CTTACTGATATGCAATGAGCAACCTTCACTTTTACTGTTTAAAGAGCGGTCAATT 672
QY 727 TTGCTGATGAATGGAATGCAATATACATCTTCAAAATTTGAACCTTAATCTGGAACA 786
DB 673 TTTGAGAGAAATGGGGAATGTCAACAAT-----ACT 705
QY 787 TCAGATGACTATTATAAATTTTAAAGAAATATTAATTAATAGTAACTATGTA 846
DB 706 ATTAATTAATTAATTAATGATGTCAAATGAAACTTACTGCAATTAATTTGATCACTGTGA 765
QY 847 AATACCTATAGAACAGACTAATAAATCTTAGAGCAACCAATATGAAATGAGATTA 906

```

```
Db 766 AAGGTGATGAACTGGTTTGGCAAAATTAAGGACAGCGCTAAACAAATGGGTTAC 825
Qy 907 TTTAATGACTATCGAAGATATATGACCACTTACTGATTAGATACCATCTCAATTTCT 966
Db 826 TATAACCAATTCGGTAGAAGAAATGACACTGGCGGTTTATGATGTTGTCATATTTCCCA 885
Qy 967 TTAATGATATTAATAAGATA 986
Db 886 AATTATGACACGACGCGTA 905
```

## RESULT 5

```
US-11-108-389-51
; Sequence 51, Application US/11108389
; Publication No. US20050261188A1
; GENERAL INFORMATION:
; APPLICANT: Andre R. Abad
; APPLICANT: Ronald D. Flannagan
; APPLICANT: Rafael Herrmann
; APPLICANT: Theodore W. Kahn
; APPLICANT: Albert L. Lu
; APPLICANT: Billy Fred McCutchen
; APPLICANT: James K. Prenail
; APPLICANT: James F.H. Wong
; APPLICANT: Cao-Guo Yu
; TITLE OF INVENTION: Genes Encoding Proteins With Pesticidal
; FILE REFERENCE: 35718/291049
; CURRENT APPLICATION NUMBER: US/11/108,389
; PRIOR FILING DATE: 2005-04-18
; PRIOR APPLICATION NUMBER: 60/391,786
; PRIOR FILING DATE: 2002-06-26
; PRIOR APPLICATION NUMBER: 60/460,787
; PRIOR FILING DATE: 2003-04-04
; PRIOR APPLICATION NUMBER: 10/606,320
; PRIOR FILING DATE: 2003-06-25
; NUMBER OF SEQ ID NOS: 134
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 51
; LENGTH: 2031
; TYPE: DNA
; ORGANISM: Bacillus thuringiensis (mutated)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(2031)
US-11-108-389-51
```

Query Match 2.8%; Score 60.8; DB 7; Length 2031;

Best Local Similarity 47.1%; Pred. No. 0.014; Matches 292; Conservative 0; Mismatches 292; Indels 36; Gaps 2;

```
Qy 367 GACAAAGCAGTATGACACAAATTTTAAATGGAGAAATTTTGTGATACCGCTTA 426
Db 322 GAAAAGATGATGGAAATTTTATGAAACAAGTAAGAACTATTAATCAAAAAATA 381
Qy 427 ACAGAAAGCATTAACACAGCTAAAGTTACAACTTTAGAGATTTAGCAAAATATTACA 486
Db 382 GCAGAAATATGCAAGAAATTAAGCGCTTCGAAATTAGAAAGTTAGTAATTAACCAA 441
Qy 487 AGCTATTAATACGACTATGATGATTTGGAGAAAATTAAGAACTCAAGCTCTCGATTA 546
Db 442 TTAATCTCAACCTGCGCTTGAAATAATGGAAATAATTCATTCGAAGTCGAGTAGTTTA 501
Qy 547 CCACCATATATGACATTAACAACAGCTGCTGACTTTAAATAATGATTTGAGAAATGT 606
Db 502 AATGATTCGGCGCAGCGCTTACGAATG-----TGCGAATCGAATTTGAAATTCCTG 552
Qy 607 CACAATGATTTTATTCGAGAAATACCTGTTTCAACTGGAATTTAATAACGCTATTA 666
Db 553 GATAGCTTATTTAAGCAATATATGCAATCTTTTAAGAGTGAACAAATTTGAGAGTACATTC 612
Qy 667 CTACCTATTTATGCGCAAGCTGCTAATTTTCATTTAAATTTATTAACAACAGGTCTGA 726
```

```
Db 613 CTACTGTATATATGCAATATGACAGCCACCTTCATTTACTGTATTAAGAGACGCTCAATT 672
Qy 727 TTGGCTGATGAATGGAATGAGATATATACCTTCCTTCACAAAATGAACTAATGTGGAACA 786
Db 673 TTGGAGAAAGATGGGATGCTCAACACT-----ACT 705
Qy 787 TCAATGACTATTTATTAACCTTTTAAAGAAATATATACCTTAATATATGATTAATTTGCA 846
Db 706 ATTAATTAATCTATATATATGATCGCAATATGAACCTTACTCAGAAATATTCGATCACTGTGTA 765
Qy 847 AATACCTATGAACAGGACTTAAATAATCTTAGAGACGACCAAAATATGAATGAGATATA 906
Db 766 AAGGTGATGAACTGGTTTGGCAAAATTAAGGACAGCGCTAAACAAATGGGTTGAC 825
Qy 907 TTTAATGACTATCGAAGATATATGACCACTTACTGATTAGATACCATCTCAATTTTCT 966
Db 826 TATAACCAATTCGGTAGAAGAAATGACACTGGCGGTTTATGATGTTGTCATATTTCCCA 885
Qy 967 TTAATGATATTAATAAGATA 986
Db 886 AATTATGACACGACGCGTA 905
```

## RESULT 6

```
US-11-108-389-83
; Sequence 83, Application US/11108389
; Publication No. US20050261188A1
; GENERAL INFORMATION:
; APPLICANT: Andre R. Abad
; APPLICANT: Ronald D. Flannagan
; APPLICANT: Rafael Herrmann
; APPLICANT: Theodore W. Kahn
; APPLICANT: Albert L. Lu
; APPLICANT: Billy Fred McCutchen
; APPLICANT: James K. Prenail
; APPLICANT: James F.H. Wong
; APPLICANT: Cao-Guo Yu
; TITLE OF INVENTION: Genes Encoding Proteins With Pesticidal
; FILE REFERENCE: 35718/291049
; CURRENT APPLICATION NUMBER: US/11/108,389
; PRIOR FILING DATE: 2005-04-18
; PRIOR APPLICATION NUMBER: 60/391,786
; PRIOR FILING DATE: 2002-06-26
; PRIOR APPLICATION NUMBER: 60/460,787
; PRIOR FILING DATE: 2003-04-04
; PRIOR APPLICATION NUMBER: 10/606,320
; PRIOR FILING DATE: 2003-06-25
; NUMBER OF SEQ ID NOS: 134
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 83
; LENGTH: 2031
; TYPE: DNA
; ORGANISM: Bacillus thuringiensis
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(2031)
US-11-108-389-83
```

Query Match 2.8%; Score 60.8; DB 7; Length 2031;

Best Local Similarity 47.1%; Pred. No. 0.014; Matches 292; Conservative 0; Mismatches 292; Indels 36; Gaps 2;

```
Qy 367 GACAAAGCAGTATGACACAAATTTTAAATGGAGAAATTTTGTGATACCGCTTA 426
Db 322 GAAAAGATGATGGAAATTTTATGAAACAAGTAAGAACTATTAATCAAAAAATA 381
Qy 427 ACAGAAAGCATTAACACAGCTAAAGTTACAACTTTAGAGATTTAGCAAAATATTACA 486
Db 382 GCAGAAATATGCAAGAAATTAAGCGCTTCGAAATTAGAAAGTTAGTAATTAACCAA 441
Qy 487 AGCTATTAATACGACTATGATGATTTGGAGAAAATTAAGAACTCAAGCTCTCGATTA 546
```



Db	442	TTATATCTAACTGCGCTTGAGAAATGGGAAGAAATTCATTTGGAAGTCGAGTACTTTA	501
Qy	547	CCACCATCATGACATTTACAACAAGCTGCTTGACTCTTAAATVACGATTTGGAATGTT	606
Db	502	AATGTTCCCGGCGACGCTTACGAGATG-----TGGAAATCGATTTGAAATCCTG	552
Qy	607	CACATGATTTTATTCGAGAAATACCTGGTTTCCACTTGAACCTTATAAACGCTATT	666
Db	553	GATGTTTATTTACGCAATATATATGCTCTTTAGAGTACAAATTTTGAAGTACATTTC	612
Qy	667	CTACCTATTATATGCGCAAGCTGCTATTTTCAATTTTAAATTTATTAACAAGGTGCTGAA	726
Db	613	CTTACTGATATATGCAATGGCAGCCAACTTCAATTTACTGTATTTAAAGACGGTCAATT	672
Qy	727	TTGGCTGATGAATGGAATGCAGATATACATCTTCAAAATTTGAACCTTAATGCTGGAACA	786
Db	673	TTTGGAGAAGGATGGGGATGTCACAACT-----ACT	705
Qy	787	TCAGATGACTATTTAACTTTTAAAGAAATATACCTAAATATAGTAACTATTGGCA	846
Db	706	ATTATTAATCTATTATGATGTCATAATGAAACTTACTGCAGAAATATTCGTGATCCTGTGA	765
Qy	847	AATACCTATTAGAACAGACTAAATAATCTTAGAGCAAGCAACAATATGAATGAGATATA	906
Db	766	AAGTGTATGAAACCTGTTTAGCAAAATTTAAAGGACGAGCGCTTAAACAATGGTTGAC	825
Qy	907	TTTAAATGACTATCGAAGATATATGACCAATTACTGTATTTAGATACCAATCTCTCAATTTCT	966
Db	826	TATTAACCAATTCCTTAGAGAAATGACACTGGCGGTTTATGATGTGTGCAATTAATCCCA	885
Qy	967	TTATATGATATTAATAAGATA	986
Db	886	AATATGACACACGACGTA	905

```

RESULT 7
US-11-121-086-25
; Sequence 25, Application US/11121086
; Publication No. US20050266459A1
; GENERAL INFORMATION:
; APPLICANT: POULSEN, TIM S.
; APPLICANT: NIELSEN, KIRSTEN V.
; TITLE OF INVENTION: NUCLEIC ACID PROBES AND NUCLEIC ACID ANALOG PROBES
; FILE REFERENCE: 09138.6000-00000
; CURRENT APPLICATION NUMBER: US/11/121, 086
; PRIOR FILING DATE: 2005-05-04
; PRIOR APPLICATION NUMBER: 60/567, 570
; PRIOR FILING DATE: 2004-05-04
; NUMBER OF SEQ ID NOS: 107
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 25
; LENGTH: 173602
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-121-086-25

```

	Query Match	2.7%	Score 60.6	DB 7	Length 173602
	Best Local Similarity	43.3%	Pred. No. 0.04		
	Matches 487	Conservative	0	Mismatches 629	Indels 9
				Gaps 4	
Qy	594	ATTGAGAGATGTCACAAATGATTTTATTTGAGAAAATACCTGGTTTCCAATTGAACTTA	653		
Db	136136	ATATATAATATATATATATATATTTATATGATATATATAATATAATATATATATATATA	136195		
Qy	654	TAAAACGCTATTACTTATTATGCGCAGCTGCTAAATTTTCATTTAAATTTATTACA	713		
Db	136196	TAAATATATTAATATATATATTAATTAATATATATTAATTAATTAATTAATTAATATA	136255		
Qy	714	ACAAAGTCCTGAATGGCT-GATGAATGGAATGCGATATACCTCTTCGCAAAATTAAC	772		
Db	136256	TAAATATATATAATATATATATATATATATAATATAATATAATATAATATAATATA	136315		

[illegible]

RESULT 8  
US-11-056-727-39  
; Sequence 39, Application US/11058722-39  
; Publication No. US20050261483A1  
; GENERAL INFORMATION:  
; APPLICANT: Andre R. Abad  
; APPLICANT: Ronald D. Pinnagan  
; APPLICANT: Rafael Herrmann

APPLICANT: Theodore W. Kahn  
APPLICANT: Albert L. Lu  
APPLICANT: Billy Fred McCutchen  
APPLICANT: James K. Presnail  
APPLICANT: James F.H. Wong  
APPLICANT: Cao-Guo Yu  
TITLE OF INVENTION: Genes Encoding Proteins With Peestcidai  
TITLE OF INVENTION: Activity  
FILE REFERENCE: 35718/287809  
CURRENT FILING DATE: 2005-02-15  
PRIOR APPLICATION NUMBER: 60/391,786  
PRIOR FILING DATE: 2002-06-26  
PRIOR APPLICATION NUMBER: 60/460,787  
PRIOR FILING DATE: 2003-04-04  
PRIOR APPLICATION NUMBER: 10/606,320  
PRIOR FILING DATE: 2003-06-25  
NUMBER OF SEQ ID NOS: 134  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 39  
LENGTH: 2028  
TYPE: DNA  
ORGANISM: Bacillus thuringiensis (mutated)  
FEATURES:  
NAME/KEY: CDS  
LOCATION: (1)...(2028)  
US-11-058-727-39

Query Match 2.7%; Score 59.6; DB 7; Length 2028;  
Best Local Similarity 47.1%; Pred. No. 0.023;

Matches 292; Conservative 0; Mismatches 289; Indels 39; Gaps 2;

367 GACAAAACAGTATGACACAAATTTTAAATGGAGAAATTTTGTGTATACCGTTA 426  
322 GAAAAGCTCATGGGAAATTTTATGGAACAAGTAAAGAACTCATTAACAAAAATA 381  
427 ACAGAAAGCATTAACAGCTAAAGTTACAACTTTAGAGAGTTTACCAATATTACA 486  
382 GCAGAAATATGCAAGAAATTAAGCGCTTTCGAAATTGAGAAATTAATTAATCA 441  
487 AGCTATATATACAGCTATGATGATGAGAAATTTAAAGACTACAGCTCCGTGATTA 546  
442 TTATATCTAATCGCTGCTTGAAGATGGAGAAATTCATTCCAGTCGAGTTTGA 501  
547 CCACCATCATGACATTAACAACAGCTGCTGACTCTTAAATACGATTGAGAAATGT 606  
502 AGTGAAGGTCCAGCTTACGAGA-----TGTGCAAAATCGATTGAAATCCTG 549  
607 CACATATGTTTATTCGAGAAATACCTGTTCCAACTTGAAACTTATTAACGCTATTA 666  
550 GATAGTTTATTTACGCAATATATGCAATCTTTAGAGTACAAATTTGAAATGACATTC 609  
667 CTACCTATTTATGCGCAAGCTGCTAATTTTCAATTTAATTTATTAACAAGGTGGA 726  
610 CTTACTGATATGCAATGAGCGCAACCTTCACTTATCTGTTATTTAAAGAGCGCTCAATT 669  
727 TTGCGTATGATGAGATGAGATGAGATATACATCTTCACAAAATGGAACCTTAATGCTGGAACA 786  
670 TTGGAAGAAATGGAGATGATGCAACACT-----ACT 702  
787 TCAGATGATCTATTAATCTTTAAAGAAATATATCTTAATATAGTAATCTTATGCA 846  
703 ATTAATACTATTAATGATGCTCAATGAACTTACTGAGAAATTTCTGATCAGCTGCTA 762  
847 AATACCTATTAAGACAGACTAATAAATCTTGAAGACGAACCAATATGAATGAGATTA 906  
763 AAGTGTATGAAATCGTTTACCAAAATTTAAAGCAGAGCGCTTAACAAATGGCTGAC 822  
907 TTTAATGATATGAGATATATATGACATTAATCTGATTAATGATACATCTCAATTTTCT 966  
823 TATTAACCAATTCGTAAGAGAAATGACACTGGCGGTTTAAAGTGTGATTAATTTCCA 882  
967 TTAATATGATTAATAAAGATA 986

Db 883 AATTATGACACAGCAGCTA 902

RESULT 9  
US-11-058-727-71

Sequence 71, Application US/11058727  
Publication No. US20050261483A1

GENERAL INFORMATION:

APPLICANT: Andre R. Abad

APPLICANT: Ronald D. Flammagan

APPLICANT: Rafael Herrmann

APPLICANT: Theodore W. Kahn

APPLICANT: Albert L. Lu

APPLICANT: Billy Fred McCutchen

APPLICANT: James K. Presnail

APPLICANT: James F.H. Wong

APPLICANT: Cao-Guo Yu

TITLE OF INVENTION: Genes Encoding Proteins With Peestcidai

TITLE OF INVENTION: Activity

FILE REFERENCE: 35718/287809

CURRENT FILING DATE: 2005-02-15

PRIOR APPLICATION NUMBER: 60/391,786

PRIOR FILING DATE: 2002-06-26

PRIOR APPLICATION NUMBER: 60/460,787

PRIOR FILING DATE: 2003-04-04

PRIOR APPLICATION NUMBER: 10/606,320

PRIOR FILING DATE: 2003-06-25

NUMBER OF SEQ ID NOS: 134

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 71

LENGTH: 2028

TYPE: DNA

ORGANISM: Bacillus thuringiensis (mutated)

FEATURES:

NAME/KEY: CDS

LOCATION: (1)...(2028)

US-11-058-727-71

Query Match 2.7%; Score 59.6; DB 7; Length 2028;  
Best Local Similarity 47.1%; Pred. No. 0.023;

Matches 292; Conservative 0; Mismatches 289; Indels 39; Gaps 2;

367 GACAAAACAGTATGACACAAATTTTAAATGGAGAAATTTTGTGTATACCGTTA 426  
322 GAAAAGCTCATGGGAAATTTTATGGAACAAGTAAAGAACTCATTAACAAAAATA 381  
427 ACAGAAAGCATTAACAGCTAAAGTTACAACTTTAGAGAGTTTACCAATATTACA 486  
382 GCAGAAATATGCAAGAAATTAAGCGCTTTCGAAATTGAGAAATTAATTAATCA 441  
487 AGCTATATATACAGCTATGATGATGAGAAATTTAAAGACTACAGCTCCGTGATTA 546  
442 TTATATCTAATCGCTGCTTGAAGATGGAGAAATTCATTCCAGTCGAGTTTGA 501  
547 CCACCATCATGACATTAACAACAGCTGCTGACTCTTAAATACGATTGAGAAATGT 606  
502 AGTGAAGGTCCAGCTTACGAGA-----TGTGCAAAATCGATTGAAATCCTG 549  
607 CACATATGTTTATTCGAGAAATACCTGTTCCAACTTGAAACTTATTAACGCTATTA 666  
550 GATAGTTTATTTACGCAATATATGCAATCTTTAGAGTACAAATTTGAAATGACATTC 609  
667 CTACCTATTTATGCGCAAGCTGCTAATTTTCAATTTAATTTATTAACAAGGTGGA 726  
610 CTTACTGATATGCAATGAGCGCAACCTTCACTTATCTGTTATTTAAAGAGCGCTCAATT 669  
727 TTGCGTATGATGAGATGAGATGAGATATACATCTTCACAAAATGGAACCTTAATGCTGGAACA 786  
670 TTGGAAGAAATGGAGATGATGCAACACT-----ACT 702  
787 TCAGATGATCTATTAATCTTTAAAGAAATATATCTTAATATAGTAATCTTATGCA 846

D <sub>b</sub>	703	ATTAAATCAATTATGATGCTCAATGAAACTTACTGAGAAATATTCGATCACTGGTA	762
Q <sub>y</sub>	847	AATACCTATAGAACAGACTAAAAAATCTAGAGACGAACAAATATGAATGAGATATA	906
D <sub>b</sub>	763	AGGTGATAGAAATCTGTTTACGAAATTTAAAGCAGACAGCGCTAAACAAATGGTTGAC	822
Q <sub>y</sub>	907	TTTAAATGACTATCGAAGATATATGACCATCTGATTAGATATACATCTCCTCAATTTCT	966
D <sub>b</sub>	823	TATTAACCAATTCGATGAGAAATGACACTGGCGGTTTAAATGTTGTCATTAATCCCA	882
Q <sub>y</sub>	967	TTTATGATATATAAAGATA	986
D <sub>b</sub>	883	AATTATGACACGACGCTTA	902

```

RESULT 10
US-11-108-389-39
; Sequence 39. Application US/11108389
; Publication No. US20050261188A1
GENERAL INFORMATION:
APPLICANT: Andre R. Abad
APPLICANT: Ronald D. Flannagan
APPLICANT: Rafael Herrmann
APPLICANT: Theodore W. Kahn
APPLICANT: Albert L. Iu
APPLICANT: Billy Fred McCutchen
APPLICANT: James K. Presnall
APPLICANT: James F.H. Wong
APPLICANT: Cao-Guo Yu
TITLE OF INVENTION: Genes Encoding Proteins With Pesticidal
; TITLE OF INVENTION: Activity
FILE REFERENCE: 35716/291049
CURRENT APPLICATION NUMBER: US/11/108,389
CURRENT FILING DATE: 2005-04-18
PRIOR APPLICATION NUMBER: 60/391,786
PRIOR FILING DATE: 2002-06-26
PRIOR APPLICATION NUMBER: 60/460,787
PRIOR FILING DATE: 2003-04-04
PRIOR APPLICATION NUMBER: 10/606,320
PRIOR FILING DATE: 2003-06-25
NUMBER OF SEQ ID NOS: 134
SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 39
; LENGTH: 2028
; TYPE: DNA
; ORGANISM: Bacillus thuringiensis (mutated)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1) ... (2028)
; US-11-108-389-39

```

Query Match	2.7%	Score	59.6	DB	7	Length	208
Best Local Similarity	47.1%	Pred. No.	0.023				
Matches	292	Conservative	0	Mismatch	289	Indels	39
						Gaps	2

Qy	367	GACAAACAGTATGGACGCAATTTATTAATATGGAGAAATTTTGTGTATACCGCTTA	426
Db	322	GAAAGAGTCATGGGAATTTTTATATGGACAAGTAGAAGACCTCAATTAATCAAAAAATA	381
Qy	427	ACGAAAGCATTAACCGCTAAAGTTACAACTTTAGAGAGATTAGACAATATTAACA	486
Db	382	GCAGATATGACAGGAATAAAGCGCTTTCCGAAATTAGAGAGATTAGTAATTAATTACA	441
Qy	487	AGCTAATATACGACATTGATGATTTGGAGAAATTAATAAGACTAACAGCTCCGTGATTA	546
Db	442	TTATATCTTACGCGCTTGAGAAATGGGAAGAAATTCATTTGCAAGTCGAGGTTTTGCA	501
Qy	547	CCACCATCATGCAATTAACAACAGGCTCCTTGACTCTTAATAATACCATTTGGAATGTT	606
Db	502	AGTCGAGGTCAGGCTTAGAGA-----TGCGCAATTCGATTTGGAATCTCG	549
Qy	607	CACAAATGATTTTATTCGAGAAATACCTGGTTTCCACTTGAACTTATTAACGCTATTA	666

Db	550	GAATGTTTATTTACGCAATATATGCGAATCTTTTAGAGTAGCAAAATTTTGAAGTACCAATTC	6090
Qy	667	CTACCTATTTATGCGCAGCTCTATTTTTCATTTAAATTTATTTACAAAGAGTGTGAA	7260
Db	610	CTTACTGTATATGCAATGCGACCACTTCACTTACTGTATTTAAAGACGCGTCAATT	6680
Qy	727	TTGCGTATGAAATGGAATGCAAGTATACCTCTTCACAAATGGAACCTAATGCTGGAACA	7860
Db	670	TTTGGAGAAGATGGGAGATGTCAACAAC-----ACT 7020	
Qy	787	TCAGATGACTATTTAACTTTTAAAGAAAAATATACCTAAATATAGTAACTATTTGCA	8460
Db	703	ATTAATPACTATTTATGATGTGCAATGAAACTTACTGCAAGATTTTGTGATCACTGTGTA	7520
Qy	847	AATACTTATGAAACGAGACTAAAAATTTTGAGACGAAACCAATATGAAATGAGATATA	9060
Db	763	AAGTGTATGAAACGTGTTTACCAAAATTTAAAGCGCAGACGCTTAAACAACGTGGTTGAC	8220
Qy	907	TTTATGACTATGGAAGATATATGCAATTACTGTATTAATATACCACTCTCAATTTTCT	9660
Db	823	TATTAACCAATTCGATAGGAATGCACTGGCGGTTTTAGATGTGTGCACTTATTTCCCA	8820
Qy	967	TTATATGATATATAAAGATA 986	
Db	883	AATTATGACACGCACTGTA 902	

RESULT 11  
US-11-108-389-71  
; Sequence 71. Application US/11108389  
; Publication No. US20050261188A1  
; GENERAL INFORMATION:  
; APPLICANT: Andre R. Abad  
; APPLICANT: Ronald D. Flannagan  
; APPLICANT: Rafael Herrmann  
; APPLICANT: Theodore W. Kahn  
; APPLICANT: Albert L. Iu  
; APPLICANT: Billy Fred McCutchen  
; APPLICANT: James K. Prenatal1  
; APPLICANT: James F.H. Wong  
; APPLICANT: Cao-Guo Yu  
; TITLE OF INVENTION: Gene Encoding Proteins With Peptidaly  
; FILE REFERENCE: 35718/291049  
; CURRENT FILING DATE: 2005-04-18  
; PRIOR APPLICATION NUMBER: 60/391,786  
; PRIOR FILING DATE: 2002-06-26  
; PRIOR APPLICATION NUMBER: 60/460,787  
; PRIOR FILING DATE: 2003-04-04  
; PRIOR APPLICATION NUMBER: 10/606,320  
; PRIOR FILING DATE: 2003-06-25  
; NUMBER OF SEQ ID NOS: 134  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 71  
; LENGTH: 2028  
; TYPE: DNA  
; ORGANISM: *Bacillus thuringiensis* (mutated)  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1) ... (2028)  
; US-11-108-389-71

Query Match	2.7%	Score 59.6	DB 7	Length 208
Best Local Similarity	47.1%	Pred. No. 0.023		
Matches 292	Conservative	0	Mismatches 289	Indels 39
				Gaps 2

**Oy** 367 GACAAAACGATGGACAATTATTAAATGGGAATAATTTTGTGTGTAACCCTTA 426  
||| ||| ||| ||| ||| ||| ||| |||  
**Db** 322 GAAAAAGCTCAATGGGAATTTTATGGACACGTAAGAACACTCATTAATCAAAAAATA 381  
||| ||| ||| ||| ||| ||| ||| |||  
**Oy** 427 ACAGAAGCATAAACAGCTAAAGTCMAAATTAAAGAAATTTTAGCAAAATATTACAA 486

```
Db 382 GCAGAAATATGCAAGAAATAAAGCGCTTTCGAAATTAGAGATTAAGTAATAATTACCA 441
Qy 487 AGCTATATATACAGATTAGATGATTTGGAGAAAATTAAGAACTACAAAGCTCTGATTA 546
Db 442 TTATATCTAACTGGCGCTTGAAGAAATGGGAGAAAATCAATTCCAGTCGAGCTTTGCA 501
Qy 547 CCACCATCATGACATTAACAACAAGCTGCTTGAAGCTTTAAATTCGATTTGAGAAATGTT 606
Db 502 AGTCAGAGTCACGCTTACGAGA-----TGTGCGAAAATCGATTTGAATCTCTG 549
Qy 607 CACAATGATTTATTCGAGAAATACCTGCTTCCAACTGAACTTATTAAGCTATTA 666
Db 550 GATGCTTATTTATTCGCAATATATGCAATCTTTTGAAGTACAAATTTGAAGTACCAATTC 609
Qy 667 CTACCTATTTATTCGCAAGCTGCTTAACTTATTTATTTATTTATTAACAACAAGTCTGAA 726
Db 610 CTTACTGATATGCAATGGCAGCCAACTTATTTATTTATTTATTTAAAGAGCGCTCAAT 669
Qy 727 TTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 786
Db 670 TTTGAGAGAGATGGGATGGTCAACACT-----ACT 702
Qy 787 TCAGATGATCTTATTAATCTTTTAAAGAAATATTAATTAATTAATTAATTAATTAATTA 846
Db 703 ATTAATTAATCTATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 762
Qy 847 AATACCTATGAGACAGCACTTAAATTTAGAGACGAAACCAATATATGAGATTA 906
Db 763 AAGTGATATGAACTGCTTTAGCAAAATTAAGGACGAGCGTAAACATGGGTTGAC 822
Qy 907 TTTAATGATCTATGAGATATATGATGATGATGATGATGATGATGATGATGATGATGATG 966
Db 823 TATAACCAATTCCTAGAGAAATGACACTGCGGTTTATGATGATGATGATGATGATGATG 882
Qy 967 TTATATGATATTAATAAGATA 986
Db 883 AATTATGACACGACGACGTA 902

RESULT 12
US-11-058-727-13
; Sequence 13, Application US/11058727
; Publication No. US20050261483A1
; GENERAL INFORMATION:
; APPLICANT: Andre R. Abad
; APPLICANT: Ronald D. Flannagan
; APPLICANT: Rafael Herrmann
; APPLICANT: Theodore W. Kahn
; APPLICANT: Albert L. Lu
; APPLICANT: Billy Fred McCutchen
; APPLICANT: James K. Presnall
; APPLICANT: James F.H. Wong
; APPLICANT: Cao-Guo Yu
; TITLE OF INVENTION: Genes Encoding Proteins With Pesticidal
; TITLE OF INVENTION: Activity
; FILE REFERENCE: 35718/287809
; CURRENT APPLICATION NUMBER: US/11/058,727
; PRIOR FILING DATE: 2005-02-15
; PRIOR APPLICATION NUMBER: 60/391,786
; PRIOR FILING DATE: 2002-06-26
; PRIOR APPLICATION NUMBER: 60/460,787
; PRIOR FILING DATE: 2003-04-04
; PRIOR APPLICATION NUMBER: 10/606,320
; PRIOR FILING DATE: 2003-06-25
; NUMBER OF SEQ ID NOS: 134
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13
; LENGTH: 2022
; TYPE: DNA
; ORGANISM: Bacillus thuringiensis (truncated)
; FEATURE:
; NAME/KEY: CDS
```

```
; LOCATION: (1)...(2022)
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (0)...(0)
; OTHER INFORMATION: 1218-2A
US-11-058-727-13

Query Match 2.6%; Score 58.2; DB 7; Length 2022;
Best Local Similarity 49.9%; Pred. No. 0.041;
Matches 199; Conservative 0; Mismatches 173; Indels 27; Gaps 1;

Qy 588 AATACGATTTGAGAAATGTTGACAAATGATTTTATTCGAGAAATACCTGTTCCAACTGA 647
Db 510 AATACGATTTGAGAAATGTTGACAAATGATTTTATTCGAGAAATACCTGTTCCAACTGA 569
Qy 648 AACTTATTAACGCTATTTACTATTTATTTATTTATTTATTTATTTATTTATTTATTTAT 707
Db 570 AATTTGAAAGTACATCTCTTACAGATATTAACAGAGGCAACCTTATTTATTTATTTAT 629
Qy 708 ATTACAAACAAGTGTGAAATGCTGATGATGATGATGATGATGATGATGATGATGATGAT 767
Db 630 ATTAAAGACGCTTCAATTTTGGAGAAATGG----- 663
Qy 768 TGAACCTAATGCTGGAACATGATGATGATGATGATGATGATGATGATGATGATGATGAT 827
Db 664 -GATGCTTACCAACCACTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 722
Qy 828 ATATGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 887
Db 723 ATATTCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 782
Qy 888 AATATGAAATGAGATATTTAATGATGATGATGATGATGATGATGATGATGATGATGAT 947
Db 783 CGCTAAACAATGGGTGACATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 842
Qy 948 TACATCTCTCAATTTTCTTATATGATATTAATAAGATA 986
Db 843 TGTGTTGATTTATTCCTCAATTAATGACACGACGTA 881

RESULT 13
US-11-108-389-13
; Sequence 13, Application US/11108389
; Publication No. US20050261188A1
; GENERAL INFORMATION:
; APPLICANT: Andre R. Abad
; APPLICANT: Ronald D. Flannagan
; APPLICANT: Rafael Herrmann
; APPLICANT: Theodore W. Kahn
; APPLICANT: Albert L. Lu
; APPLICANT: Billy Fred McCutchen
; APPLICANT: James K. Presnall
; APPLICANT: James F.H. Wong
; APPLICANT: Cao-Guo Yu
; TITLE OF INVENTION: Genes Encoding Proteins With Pesticidal
; TITLE OF INVENTION: Activity
; FILE REFERENCE: 35718/291049
; CURRENT APPLICATION NUMBER: US/11/108,389
; PRIOR FILING DATE: 2005-04-18
; PRIOR APPLICATION NUMBER: 60/391,786
; PRIOR FILING DATE: 2002-06-26
; PRIOR APPLICATION NUMBER: 60/460,787
; PRIOR FILING DATE: 2003-04-04
; PRIOR APPLICATION NUMBER: 10/606,320
; PRIOR FILING DATE: 2003-06-25
; NUMBER OF SEQ ID NOS: 134
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13
; LENGTH: 2022
; TYPE: DNA
; ORGANISM: Bacillus thuringiensis (truncated)
; FEATURE:
; NAME/KEY: CDS
```

LOCATION: (1)...(2022)  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (0)...(0)  
OTHER INFORMATION: 1218-2A  
US-11-108-389-13

Query Match 2.6%; Score 58.2; DB 7; Length 2022;  
Best Local Similarity 49.9%; Pred. No. 0.041;  
Matches 199; Conservative 0; Mismatches 173; Indels 27; Gaps 1;

588 AATACGATTTGAGAAATGTTCAATGATTTTATTCGAGAAATACCTGGTTCCAACTTGA 647  
510 AATCGATTTGAAATCCGATGATGTTTATTTACGCAATACATGCGATCTTTTCGAGTAC 569  
648 AACTATATAACGCTATACCTATTTATGCGCAAGCTGTATTTTCAATTTAAATTT 707  
570 AATTTTGAAGTACCATTCCTTACAGTATATACAGAGCAGCCAACTTCAATTTACTGTT 629  
708 ATTACAAACAGGTGCTGAATGCTGATGAAATGCAATGCAATATACATCTTCACAAT 767  
630 ATTAAAGGACGCTTCAATTTTGGAGAAAGATG----- 663  
768 TGAACCTAATGCTGGAACATGATGACTATTTATAAATCTTTAAAGAAATATACCTPA 827  
664 -GATGCTCTCAACACACTATTTAATTAATGATGATGCTCAATGAAATGAACTTACGAGA 722  
828 ATATAGTACTATTTGCAAAATACCTATATGAAACAGAGCTAAATAATCTTAAAGAGAAC 887  
723 ATATTTGATACCTGTGTAAAGTGTATGAACTGTTTACAAATTTAAAGGACGAG 782  
888 AATATGAAATGAGATATATTAATGACTATGCAATGAAATATGACCATTAATGATTTAGA 947  
783 CGCTAAACATGGGTGCACTATTAACCAATTCCTGTAGAGAAATGACACTGACGTTTAGA 842  
948 TACCATCTCTCAATTTCTTTATATGATATATAAAGATA 986  
843 TGTGTGCTATTTATCCCAATTTATGACACGACGTA 881

RESULT 14  
US-11-058-727-3  
Sequence 3, Application US/11058727  
Publication No. US20050261483A1  
GENERAL INFORMATION:  
APPLICANT: Andre R. Abad  
APPLICANT: Ronald D. Flannagan  
APPLICANT: Rafael Herrmann  
APPLICANT: Theodore W. Kahn  
APPLICANT: Albert L. Lu  
APPLICANT: Billy Fred McCutchen  
APPLICANT: James K. Presnall  
APPLICANT: James F.H. Wong  
APPLICANT: Cao-Guo Yu  
TITLE OF INVENTION: Genes Encoding Proteins With Pesticidal  
FILE REFERENCE: 35718/287809  
CURRENT APPLICATION NUMBER: US/11/058,727  
CURRENT FILING DATE: 2005-02-15  
PRIOR APPLICATION NUMBER: 60/391,786  
PRIOR FILING DATE: 2002-06-26  
PRIOR APPLICATION NUMBER: 60/460,787  
PRIOR FILING DATE: 2003-04-04  
PRIOR APPLICATION NUMBER: 10/606,320  
PRIOR FILING DATE: 2003-06-25  
NUMBER OF SEQ ID NOS: 134  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 3  
LENGTH: 3633  
TYPE: DNA  
ORGANISM: Bacillus thuringiensis  
FEATURE:  
NAME/KEY: CDS

LOCATION: (1)...(3633)  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (0)...(0)  
OTHER INFORMATION: C1218-2  
US-11-058-727-3

Query Match 2.6%; Score 58.2; DB 7; Length 3633;  
Best Local Similarity 49.9%; Pred. No. 0.046;  
Matches 199; Conservative 0; Mismatches 173; Indels 27; Gaps 1;

588 AATACGATTTGAGAAATGTTCAATGATTTTATTCGAGAAATACCTGGTTCCAACTTGA 647  
510 AATCGATTTGAAATCCGATGATGTTTATTTACGCAATACATGCGATCTTTTCGAGTAC 569  
648 AACTATATAACGCTATACCTATTTATGCGCAAGCTGTATTTTCAATTTAAATTT 707  
570 AATTTTGAAGTACCATTCCTTACAGTATATACAGAGCAGCCAACTTCAATTTACTGTT 629  
708 ATTACAAACAGGTGCTGAATGCTGATGAAATGCAATGCAATATACATCTTCACAAT 767  
630 ATTAAAGGACGCTTCAATTTTGGAGAAAGATG----- 663  
768 TGAACCTAATGCTGGAACATGATGACTATTTATAAATCTTTAAAGAAATATACCTPA 827  
664 -GATGCTCTCAACACACTATTTAATTAATGATGATGCTCAATGAAATGAACTTACGAGA 722  
828 ATATAGTACTATTTGCAAAATACCTATATGAAACAGAGCTAAATAATCTTAAAGAGAAC 887  
723 ATATTTGATACCTGTGTAAAGTGTATGAACTGTTTACAAATTTAAAGGACGAG 782  
888 AATATGAAATGAGATATATTAATGACTATGCAATGAAATATGACCATTAATGATTTAGA 947  
783 CGCTAAACATGGGTGCACTATTAACCAATTCCTGTAGAGAAATGACACTGACGTTTAGA 842  
948 TACCATCTCTCAATTTCTTTATATGATATATAAAGATA 986  
843 TGTGTGCTATTTATCCCAATTTATGACACGACGTA 881

RESULT 15  
US-11-108-389-3  
Sequence 3, Application US/11108389  
Publication No. US20050261188A1  
GENERAL INFORMATION:  
APPLICANT: Andre R. Abad  
APPLICANT: Ronald D. Flannagan  
APPLICANT: Rafael Herrmann  
APPLICANT: Theodore W. Kahn  
APPLICANT: Albert L. Lu  
APPLICANT: Billy Fred McCutchen  
APPLICANT: James K. Presnall  
APPLICANT: James F.H. Wong  
APPLICANT: Cao-Guo Yu  
TITLE OF INVENTION: Genes Encoding Proteins With Pesticidal  
FILE REFERENCE: 35718/291049  
CURRENT APPLICATION NUMBER: US/11/108,389  
CURRENT FILING DATE: 2005-04-18  
PRIOR APPLICATION NUMBER: 60/391,786  
PRIOR FILING DATE: 2002-06-26  
PRIOR APPLICATION NUMBER: 60/460,787  
PRIOR FILING DATE: 2003-04-04  
PRIOR APPLICATION NUMBER: 10/606,320  
PRIOR FILING DATE: 2003-06-25  
NUMBER OF SEQ ID NOS: 134  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 3  
LENGTH: 3633  
TYPE: DNA  
ORGANISM: Bacillus thuringiensis  
FEATURE:  
NAME/KEY: CDS

```
; LOCATION: (1)...(3633)
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (0)...(0)
; OTHER INFORMATION: CXY1218-2
US-11-108-389-3
```

```
Query Match      2.6%; Score 58.2; DB 7; Length 3633;
Best Local Similarity 49.9%; Pred. No. 0.046;
Matches 199; Conservative 0; Mismatches 173; Indels 27; Gaps 1;
```

```
QY 588 AATGAGATTGAGATGTTCAATGATTTATTCGAGAAATACCTGGTTCCAACTTGA 647
Db 510 AATGAGATTGAGATGTTCAATGATTTATTCGAGAAATACCTGGTTCCAACTTGA 647
QY 648 AACTATATAACGCTATTTACTTACCTATTTTANGCGCAAGCTGCTAATTTTCATTTAAATTT 707
Db 570 AATTTTGAAGTACCATTCCTTACAGTATATACACAGGCAAGCCAACTTCATTTACTGTT 629
QY 708 ATTACACACAGGTGCTGATTTGGCTGATGAAATGGAATGAGATATACATCCTTCACAAT 767
Db 630 ATTAAAGGACCGCTTCAATTTTGGAGAAATGG----- 663
QY 768 TGAACCTTAATGCTGGAACATGAGATGATTAATTAATTAATTAATTAATTAATTAATTA 827
Db 664 -GGAATGCTTAACAAACCACTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 722
QY 828 ATATAGTAACTATTGTGCAAAATACCTATATGAAACAGACTAATAAAATCTTAGAGACGAA 887
Db 723 ATATTTGATGACAGTGTGTAAGGTGTATGAAAGTGTATGCAAAATTAATAAGCAGAG 782
QY 888 AAATATGAAATGAGATATATTTAATGACTATGAGAGATATATGACCATTTACTGTATTAGA 947
Db 783 CGCTAAACAAATGGGTGCACTATTAACCAATTCGGTATGAGAAATGACACTGACGTTTGA 842
QY 948 TACCATCTTCATTTTCTTTATATGATATTAATAAGATA 986
Db 843 TGTGTGTGCAATTAATTCCTCAAAATTAATGACACGACGCTA 881
```

Search completed: December 11, 2005, 14:11:53  
Job time : 223 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using SW model

Run on: December 11, 2005, 03:43:39 ; Search time 1791 Seconds  
(without alignments)  
10194.744 Million cell updates/sec

Title: US-10-783-417-1

Perfect score: 2208  
Sequence: 1 atgatacaataacagataa.....atcacatacaaacacttga 2208

Scoring table: IDENTITY NUC  
Gapop 10.0, Gapext 1.0

Searched: 9793542 seqs, 4134689005 residues

Total number of hits satisfying chosen parameters: 19587084

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

Published Applications NA\_Main:  
1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq:\*  
2: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq:\*  
3: /cgn2\_6/ptodata/1/pubpna/US09A\_PUBCOMB.seq:\*  
4: /cgn2\_6/ptodata/1/pubpna/US09B\_PUBCOMB.seq:\*  
5: /cgn2\_6/ptodata/1/pubpna/US10A\_PUBCOMB.seq:\*  
6: /cgn2\_6/ptodata/1/pubpna/US10B\_PUBCOMB.seq:\*  
7: /cgn2\_6/ptodata/1/pubpna/US10C\_PUBCOMB.seq:\*  
8: /cgn2\_6/ptodata/1/pubpna/US10D\_PUBCOMB.seq:\*  
9: /cgn2\_6/ptodata/1/pubpna/US10E\_PUBCOMB.seq:\*  
10: /cgn2\_6/ptodata/1/pubpna/US11\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2208	100.0	2208	US-10-783-417-1	Sequence 1, Appli
2	1903.8	86.2	2235	US-10-782-570-1	Sequence 1, Appli
3	1794.4	81.3	2085	US-10-782-570-3	Sequence 3, Appli
4	163.8	7.4	4896	US-09-756-526A-3	Sequence 3, Appli
5	163.8	7.4	4896	US-10-345-020-3	Sequence 3, Appli
6	163.8	7.4	4896	US-10-342-821-3	Sequence 3, Appli
7	98.4	4.5	3504	US-10-089-678-2	Sequence 2, Appli
8	98.4	4.5	3690	US-10-089-678-3	Sequence 3, Appli
9	97	4.4	3673778	US-10-312-841-1	Sequence 1, Appli
C 10	94.6	4.1	15548	US-10-311-455-2128	Sequence 2128, Ap
C 11	90.4	4.1	8056	US-10-473-126-386	Sequence 386, App
12	85.8	3.9	3522	US-09-826-660-5	Sequence 5, Appli
13	85.8	3.9	3522	US-09-837-961-7	Sequence 7, Appli
14	85.8	3.9	3522	US-10-825-751-7	Sequence 7, Appli
C 15	81.2	3.7	6175	US-10-311-455-180	Sequence 1280, Ap
C 16	77.8	3.5	3673778	US-10-312-841-2	Sequence 2, Appli
C 17	76.4	3.5	3507	US-10-428-961-21	Sequence 21, Appli
C 18	74.2	3.4	8056	US-10-473-126-240	Sequence 240, App
C 19	73.2	3.3	1959	US-10-614-076-13	Sequence 13, Appli
C 20	72.6	3.3	6641	US-10-311-455-288	Sequence 288, App
C 21	72.6	3.3	6641	US-10-240-452-36	Sequence 36, Appli
C 22	72.6	3.3	8237	US-10-221-714A-527	Sequence 527, App
C 23	71.8	3.3	7498	US-10-311-455-230	Sequence 230, App

24	71.6	3.2	1959	7	US-10-614-076-27	Sequence 27, Appli
25	71.6	3.2	3684	9	US-10-929-754-2	Sequence 2, Appli
26	71	3.2	8056	8	US-10-473-126-386	Sequence 386, App
C 27	70.8	3.2	11092	6	US-10-311-455-1486	Sequence 1486, Ap
C 28	70.4	3.2	5928	6	US-10-311-455-2059	Sequence 2059, Ap
C 29	70.4	3.2	11745	6	US-10-240-453-206	Sequence 206, App
30	70	3.2	1482	7	US-10-614-076-69	Sequence 69, Appli
31	70	3.2	1956	7	US-10-614-076-51	Sequence 51, Appli
32	70	3.2	1956	7	US-10-614-076-55	Sequence 55, Appli
33	70	3.2	1956	7	US-10-614-076-57	Sequence 57, Appli
34	70	3.2	1959	5	US-10-232-665-1	Sequence 1, Appli
35	70	3.2	1959	7	US-10-614-076-1	Sequence 19, Appli
36	70	3.2	1959	7	US-10-614-076-19	Sequence 41, Appli
37	70	3.2	1959	7	US-10-614-076-41	Sequence 43, Appli
38	70	3.2	1959	7	US-10-614-076-43	Sequence 45, Appli
39	70	3.2	1959	7	US-10-614-076-45	Sequence 65, Appli
40	70	3.2	1959	7	US-10-614-076-65	Sequence 67, Appli
41	70	3.2	1959	7	US-10-614-076-67	Sequence 97, Appli
42	70	3.2	1959	7	US-10-614-076-97	Sequence 102, App
43	70	3.2	2280	7	US-10-614-076-102	Sequence 3, Appli
44	70	3.2	3507	7	US-10-614-524-3	Sequence 17, Appli
45	69.4	3.1	1959	7	US-10-614-076-17	

# ALIGNMENTS

RESULT 1  
US-10-783-417-1  
; Sequence 1, Application US/10783417  
; Publication No. US20040216186A1  
; GENERAL INFORMATION:  
; APPLICANT: Carozzi, Nadine  
; APPLICANT: Hargise, Tracy  
; APPLICANT: Koziele, Michael G.  
; APPLICANT: Duck, Nicholas B.  
; APPLICANT: Carr, Brian  
; TITLE OF INVENTION: AXMI-006, A Delta-Endotoxin Gene and  
; TITLE OF INVENTION: Methods for Its Use  
; FILE REFERENCE: 045600/274146  
; CURRENT APPLICATION NUMBER: US/10/783,417  
; CURRENT FILING DATE: 2004-02-20  
; PRIOR APPLICATION NUMBER: 60/448,806  
; PRIOR FILING DATE: 2003-02-20  
; NUMBER OF SEQ ID NOS: 15  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 1  
; LENGTH: 2208  
; TYPE: DNA  
; ORGANISM: Bacillus thuringiensis  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)...(2208)  
US-10-783-417-1

Query Match 100.0%; Score 2208; DB 8; Length 2208;  
Best local similarity 100.0%; Pred. No. 0;  
Matches 2208; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	ATGAATCAAAATTAACGATATACGAATATGAAATTTATGATTGCACTACCTTAT	60
DB	1	ATGAATCAAAATTAACGATATATACGAATATGAAATTTATGATTGCACTACCTTAT	60
QY	61	TTTCGGAACAAACAGTAATGATCTGATACCTTACACAAATTAATCCAAATCAACA	120
DB	61	TTTCGGAACAAACAGTAATGATCTGATACCTTACACAAATTAATCCAAATCAACA	120
QY	121	TTTCAAAACAAATTAACAAAGTGTCTCAATATGTCTCAAGGAATATACATATATGT	180
DB	121	TTTCAAAACAAATTAACAAAGTGTCTCAATATGTCTCAAGGAATATACATATATGT	180
QY	181	GATATATTCGAGACATTTGCTGATGCTGATACAAATGCTGCAATTAATGTCAGATAT	240

Db 181 GATAATTTGAGACATTTGCTAGTGTGCTATGCAATTTGCTGCAATTTAGTGTGCAATTTACTATT 240  
Qy 241 GTATCCGGTACTCTGTTAGCCGGTATAGTGGGCTCACTTCTATATCCGACCGATAGGA 300  
Db 241 GTATCCGGTACTCTGTTAGCCGGTATAGTGGGCTCACTTCTATATCCGACCGATAGGA 300  
Qy 301 ATATATAGGTGCTATATATATCTTTTGGTACCCCTAATCACTGTCTTTTGGCCCGGGGA 360  
Db 301 ATATATAGGTGCTATATATATCTTTTGGTACCCCTAATCACTGTCTTTTGGCCCGGGGA 360  
Qy 361 GAACAAGCAAAACAGTATGACACAAATTTATATAATGGAGAAATTTTGTGTATCA 420  
Db 361 GAACAAGCAAAACAGTATGACACAAATTTATATAATGGAGAAATTTTGTGTATCA 420  
Qy 421 CCGTTACAGAAAGCATATAACAGTTAAAGTTACAACTTTAGAGATTTAGCAATA 480  
Db 421 CCGTTACAGAAAGCATATAACAGTTAAAGTTACAACTTTAGAGATTTAGCAATA 480  
Qy 481 TTACAAGCTATATATAGCAATTTAGATGATTTGAGAAATTTTAAAGATCAAGCTCT 540  
Db 481 TTACAAGCTATATATAGCAATTTAGATGATTTGAGAAATTTTAAAGATCAAGCTCT 540  
Qy 541 GGATTTACACCATCATCGACATTTACACAGAGCTGCTTGACTCTTAAATTCGATTTGAG 600  
Db 541 GGATTTACACCATCATCGACATTTACACAGAGCTGCTTGACTCTTAAATTCGATTTGAG 600  
Qy 601 AATGTTCACATGATTTTATTCGAGAAATACCTGTGTTCAACTTGAATTTTAAACG 660  
Db 601 AATGTTCACATGATTTTATTCGAGAAATACCTGTGTTCAACTTGAATTTTAAACG 660  
Qy 661 CTATTTACCTATTTATGCGCAAGCTGCTAATTTTCAATTTTAAATTTTATTCACAGGT 720  
Db 661 CTATTTACCTATTTATGCGCAAGCTGCTAATTTTCAATTTTAAATTTTATTCACAGGT 720  
Qy 721 GCTGAATTTGCTGATGATGAGATGACAGATATACATCTTCAAAATTTGAACCTTAATCT 780  
Db 721 GCTGAATTTGCTGATGATGAGATGACAGATATACATCTTCAAAATTTGAACCTTAATCT 780  
Qy 781 GGAACATCAGATGATTTATTTAAAGAAATATATCTTAAATATATAGTACTAT 840  
Db 781 GGAACATCAGATGATTTATTTAAAGAAATATATCTTAAATATATAGTACTAT 840  
Qy 841 TGTGCAAAATACCTTATAGACAGACCTAATAATCTTAGAGCGAACCAATATGAAATGG 900  
Db 841 TGTGCAAAATACCTTATAGACAGACCTAATAATCTTAGAGCGAACCAATATGAAATGG 900  
Qy 901 AGTATATTTATGACTATCGAAGATATATGACCACTTACTGTATTTAGATACATCTTCA 960  
Db 901 AGTATATTTATGACTATCGAAGATATATGACCACTTACTGTATTTAGATACATCTTCA 960  
Qy 961 TTTTCTTATATGATATAAAGATATAGAGATTTCAATAGAGAGATAGAGTAAAGGC 1020  
Db 961 TTTTCTTATATGATATAAAGATATAGAGATTTCAATAGAGAGATAGAGTAAAGGC 1020  
Qy 1021 ATTAAGAAATGAACTCACAGAGAAATTTATACACTGAAATTAATTTTGTGCTTCT 1080  
Db 1021 ATTAAGAAATGAACTCACAGAGAAATTTATACACTGAAATTAATTTTGTGCTTCT 1080  
Qy 1081 CAACCTTAGAGTTCAACCAATCTAGCTAGATGAGATTTTAAACGCGTGAAGTTT 1140  
Db 1081 CAACCTTAGAGTTCAACCAATCTAGCTAGATGAGATTTTAAACGCGTGAAGTTT 1140  
Qy 1141 AAATTTATTTTCAATTTTGAACCAATTTATTTTATACAGAAATTAACAAATTTGGGAT 1200  
Db 1141 AAATTTATTTTCAATTTTGAACCAATTTATTTTATACAGAAATTAACAAATTTGGGAT 1200  
Qy 1201 CGTTAGTGTGATTTCTATCGTATGACACTTATAGCAATCTATATCTGAAACT 1260  
Db 1201 CGTTAGTGTGATTTCTATCGTATGACACTTATAGCAATCTATATCTGAAACT 1260  
Qy 1261 TTATATGAGAAAGAAAGAGTTGACCCGACAAACAAATPAGACCAATTTGAATCTTAT 1320  
Db 1261 TTATATGAGAAAGAAAGAGTTGACCCGACAAACAAATPAGACCAATTTGAATCTTAT 1320

Qy 1321 AAAGTTCAAATGTATCTGATAGCAATCACTCTGTTTTCCCTATTCAACACACTTT 1380  
Db 1321 AAAGTTCAAATGTATCTGATAGCAATCACTCTCTGTTTTCCCTATTCAACACACTTT 1380  
Qy 1381 ATATATATGCAATTTGAACTTTATTTAAATGGCTCATCTATCAACACACTTAAATTTCA 1440  
Db 1381 ATATATATGCAATTTGAACTTTATTTAAATGGCTCATCTATCAACACACTTAAATTTCA 1440  
Qy 1441 GCAGAGAGGTCTTATCTATATATATCAAAACAACTTTTTCATTTCTAGAAAAAA 1500  
Db 1441 GCAGAGAGGTCTTATCTATATATATCAAAACAACTTTTTCATTTCTAGAAAAAA 1500  
Qy 1501 GACTGCAATCTAGTATTTGATTCAGGTTGTTACCAAACTTTTAACTATATGATCATTT 1560  
Db 1501 GACTGCAATCTAGTATTTGATTCAGGTTGTTACCAAACTTTTAACTATATGATCATTT 1560  
Qy 1561 TTATCCATTTTTCATTTATTTACTATTTCTATGATGATTTGATTTGATTTGATTTGATTT 1620  
Db 1561 TTATCCATTTTTCATTTATTTACTATTTCTATGATGATTTGATTTGATTTGATTTGATTT 1620  
Qy 1621 GATPACAGTGTATTTAGATGAGACACAGATGATTTGATGATTTGATTTGATTTGATTTGAT 1680  
Db 1621 GATPACAGTGTATTTAGATGAGACACAGATGATTTGATGATTTGATTTGATTTGATTTGAT 1680  
Qy 1681 AAAATATTTCAATGATCCAGCAATCAAAAGTTAACAATCTTGAATCAAACTTAAAGTA 1740  
Db 1681 AAAATATTTCAATGATCCAGCAATCAAAAGTTAACAATCTTGAATCAAACTTAAAGTA 1740  
Qy 1741 ATTGAAGAGCTGTGCTATACAGAGAGAACTTGTGTTATTTACAAATCAAGGCGTTTA 1800  
Db 1741 ATTGAAGAGCTGTGCTATACAGAGAGAACTTGTGTTATTTACAAATCAAGGCGTTTA 1800  
Qy 1801 GAAATTTACATGTAATCTCTAATTTCTACACAATCTTATTTCACTTATGATCTGATATGCT 1860  
Db 1801 GAAATTTACATGTAATCTCTAATTTCTACACAATCTTATTTCACTTATGATCTGATATGCT 1860  
Qy 1861 ACAATGTGCTGGAATATCTCTCTAATATATCTCTCAAAATCAAGAGATTAATAGGA 1920  
Db 1861 ACAATGTGCTGGAATATCTCTCTAATATATCTCTCAAAATCAAGAGATTAATAGGA 1920  
Qy 1921 ATACCACTGCAAGACCTCAACCAACTTTTCTGTGATCAAAATTTAAATTTTCAATAC 1980  
Db 1921 ATACCACTGCAAGACCTCAACCAACTTTTCTGTGATCAAAATTTTAAATTTTCAATAC 1980  
Qy 1981 GGAAGATTTGGTATTTCCAAATTTCCAAATGAGTAACTTAACTTAAATTCGAAACATA 2040  
Db 1981 GGAAGATTTGGTATTTCCAAATTTCCAAATGAGTAACTTAACTTAAATTCGAAACATA 2040  
Qy 2041 CCATTTATTTTAACTGTCAGATGATTCAAATTTCAATTTTAACTGATTAATAATTTGAA 2100  
Db 2041 CCATTTATTTTAACTGTCAGATGATTCAAATTTCAATTTTAACTGATTAATAATTTGAA 2100  
Qy 2101 TTTATACCAATTTCTCTCTATAGCACCAAAATTTGAGAAAAACAAATTTGAAATCTATC 2160  
Db 2101 TTTATACCAATTTCTCTCTATAGCACCAAAATTTGAGAAAAACAAATTTGAAATCTATC 2160  
Qy 2161 CAAACAAAAATTAATACATTTTTCACAAATCATACAAAAACACTTTTGA 2208  
Db 2161 CAAACAAAAATTAATACATTTTTCACAAATCATACAAAAACACTTTTGA 2208

RESULT 2  
US-10-782-570-1  
; Sequence 1, Application US/10782570  
; Publication No. US20040210965A1  
; GENERAL INFORMATION:  
; APPLICANT: Carozzi, Nadine  
; APPLICANT: Hargiles, Tracy  
; APPLICANT: Koziele, Michael G.  
; APPLICANT: Duck, Nicholas B.  
; APPLICANT: Carr, Brian  
; TITLE OF INVENTION: AXMT-007, A Delta-Endotoxin Gene and





Db 1862 GTGCTGAGAAATCTTCTTAATATCTCTTACAAATCCAGGATATAGAAATACAC 1921  
QY 1928 CTCAACGACTCAACACACTTTTCTGATGCAAAATATATATATATATATATATATAT 1987  
Db 1922 CTCAACGACTCAACACACTTTTCTGATGCAAAATATATATATATATATATATATAT 1981  
QY 1988 TTGGGATATTTCCAAATTTCCAAATGACATGATACATTTTAAATCGAAATACATTTTA 2047  
Db 1982 TTGGGATATTTCCAAATTTCCAAATGACATGATACATTTTAAATCGAAATACATTTTA 2041  
QY 2048 TATTAATTCGCGAGATGATCAATTTTAAATGATGATGATGATGATGATGATGATGAT 2107  
Db 2042 TATTAATTCGCGAGATGATCAATTTTAAATGATGATGATGATGATGATGATGATGAT 2101  
QY 2108 CAATTAATTCCTCTATGACCAAAATAGAGAAAAAATTTGAAATCTATCCAAACAA 2167  
Db 2102 CAATTAATTCCTCTATGACCAAAATAGAGAAAAAATTTGAAATCTATCCAAACAA 2161  
QY 2168 AAATTAATTAATTAATTTTTCACAAATCATACAAAAACACTTTGA 2208  
Db 2162 AAATTAATTAATTAATTTTTCACAAATCATACAAAAATTAATTTA 2202

RESULT 3  
US-10-782-570-3  
Sequence 3, Application US/10782570  
Publication No. US20040210965A1  
GENERAL INFORMATION:  
APPLICANT: Carozzi, Nadine  
APPLICANT: Hargis, Tracy  
APPLICANT: Koziel, Michael G.  
APPLICANT: Duck, Nicholas B.  
APPLICANT: Carr, Brian  
TITLE OF INVENTION: AXMI-007, A Delta-Bndotoxin Gene and  
FILE REFERENCE: Methods for its Use  
CURRENT FILING DATE: 2004-02-19  
PRIOR FILING DATE: 2003-02-20  
NUMBER OF SEQ ID NOS: 17  
SOFTWARE: PasteSeq for Windows Version 4.0  
SEQ ID NO 3  
LENGTH: 2085  
TYPE: DNA  
ORGANISM: Bacillus thuringiensis  
FEATURES:  
NAME/KEY: CDS  
LOCATION: (1)...(2085)  
US-10-782-570-3

Query Match 81.3%; Score 1794.4; DB 8; Length 2085;  
Best Local Similarity 92.7%; Pred. No. 0;  
Matches 1908; Conservative 0; Mismatches 141; Indels 9; Gaps 2;

QY 154 ATGTGTCAAGGAATACACATATGTGATATTTTCAGACATTTGCTAGTGTATACA 213  
Db 1 ATGTGTCAAGGAATACACATATGTGATATTTTCAGACATTTGCTAGTGTATACA 60  
QY 214 ATTGTCTGAGTATGACGATCTATTTGTATCCGGTACTCTGTTAGCCGGTATAGTGG 273  
Db 61 ATTGTCTGAGTATGACGATCTATTTGTATCCGGTACTCTGTTAGCCGGTATAGTGG 120  
QY 274 CTCATCTCTATATCCGACGATAGAAATATAGTGTCTATATATATCTTTTGGTACC 333  
Db 121 CTCATCTCTATATCCGACGATAGAAATATAGTGTCTATATATATCTTTTGGTACC 180  
QY 334 CTATATCACTGCTTTTGGCCCGGAGAAACAGACAAACAGATGACACAAATTTATT 393  
Db 181 CTATATCACTGCTTTTGGCCCGGAGAAACAGACAAACAGATGACACAAATTTATT 240  
QY 394 AAAATGGAGAAATTTTGTGATACACCGTTAAACAGAAAGCATTAACAGCTAAAGTTA 453

Db 241 AAAATGGAGAAATTTTGTGATACACCGTTAAACAGAAAGCATTAACAGCTAAAGTTA 300  
QY 454 CAAACCTTAGAAGATTTAGACAAATTTTACAAAGCTATATATAGATATAGATTTGG 513  
Db 301 CAAACCTTAGAAGATTTAGACAAATTTTACAAAGCTATATATAGATATAGATTTGG 360  
QY 514 AGAAATTAAGAAAGACTACAAAGCTCTGATTAACCAATCATCATGATTAACAAGCT 573  
Db 361 AGAAATTAAGAAAGACTACAAAGCTCTGATTAACCAATCATCATGATTAACAAGCT 420  
QY 574 GCTTGAATCTTAAATATGATTTGAGATGTTTCAATGATTTTATTCGAAATATCT 633  
Db 421 GCTTGAATCTTAAATATGATTTGAGATGTTTCAATGATTTTATTCGAAATATCT 480  
QY 634 GGTTCACACTGAAATCTTAAAGCTATTTCTACTATTTATGCGCAAGCTTAAT 693  
Db 481 GGTTCACACTGAAATCTTAAAGCTATTTCTACTATTTATGCGCAAGCTTAAT 540  
QY 694 TTTCATTTAAATTTATTAACAAGGCTGAATTTGCTGATGATGATGATGATGATGATGAT 753  
Db 541 TTTCATTTAAATTTATTAACAAGGCTGAATTTGCTGATGATGATGATGATGATGATGAT 600  
QY 754 CATCTTCAAAATTTGAACCTTAATGCTGAAACATGATGATCTATTAATCTTTTAAA 813  
Db 601 CATCTTCAAAATTTGAACCTTAATGCTGAAACATGATGATCTATTAATCTTTTAAA 660  
QY 814 GAAATATATACCTAAATATAGTATCTATTTGCAAAATCTATGAAACAGACTAAAAAT 873  
Db 661 GAAATATATACCTAAATATAGTATCTATTTGCAAAATCTATGAAAGAGCTAAATTA 720  
QY 874 CTAGAGACGAAACCAATATGAATGAGATATTTATGATGATGATGATGATGATGATGAT 933  
Db 721 CTAGAGACGAAACCAATATGAATGAGATATTTATGATGATGATGATGATGATGATGAT 780  
QY 934 ATTAATGATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 993  
Db 781 ATTAATGATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 840  
QY 994 TCAATAGAGAAATGAGAAATGAAAGGCTTAAGATGAACTCAACAGAAATTTATACA 1053  
Db 841 TCAATAGAGAAATGAGAAATGAAAGGCTTAAGATGAACTCAACAGAAATTTATACA 894  
QY 1054 ACTGAAATTAATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1113  
Db 895 ACTGAAATTAATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 954  
QY 1114 GAATATATTTTACAGGTGAAATTTATTTTATTTTATTTTATTTTATTTTATTTT 1173  
Db 955 GAATATATTTTACAGGTGAAATTTATTTTATTTTATTTTATTTTATTTTATTTT 1014  
QY 1174 TATACAGAAATTAATTTTCCGGAATCGTTTATGATGATGATGATGATGATGATGATGAT 1233  
Db 1015 TATACAGAAATTAATTTTCCGGAATCGTTTATGATGATGATGATGATGATGATGATGAT 1074  
QY 1234 ACTTATAGCAATCTATTAATCTTATATGAGAAAGAAACAGGTTACACCAACA 1293  
Db 1075 ACTTATAGCAATCTTATTAATCTTATATGAGAAAGAAACAGGTTACACCAACA 1134  
QY 1294 AAAACATTAAGACATTTGAAATCTTATTAAGTTTCAATTTGATGATGATGATGATGAT 1353  
Db 1135 AAAACATTAATTAATCTTATTAAGTTTCAATTTGATGATGATGATGATGATGATGAT 1194  
QY 1354 CCGTTTCCCTTATTC--AACCACTTATATATTAATCAAAATTTGAATTTTAAAT 1410  
Db 1195 CCGTTTCCCTTATTC--AACCACTTATATATTAATCAAAATTTGAATTTTAAAT 1254  
QY 1411 GGTCAATTAACAACACTCAATATTCAGAGAGGCTTTATCTAATTAATTAATTAATTAAT 1470  
Db 1255 AATTCACTGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1314  
QY 1471 ACAATTTTCTCAATTTCTAGAAAAAAGACTGCAATCTAGTATTAATCCAGGTTGT 1530

Db 1315 ACAACGTATTTTCATTTCTCGTAAAAAAGACGTGAAACCAATATATTAATCCAAATGTG 1374

Qy 1531 TCACCAAACTTAAATPACTATAGTCATATTTTATCCCATTTTTCATTATTTACTTATCC 1590

Db 1375 TTACCAAGCTATTAATAGTATATAGTCATATTTTATCCAGTTCCTTATTTTAAATTAATCC 1434

Qy 1591 TATGTGATTTGGATTACAGCTACCAATATATTTAGATACAGGTGATTAAGATGAGACACAGT 1650

Db 1435 TATATAATTTGATTAAGCGCTAAATATATTTATTAATACAGGTGATTAAGATGAGACACAGT 1494

Qy 1651 AGTGTGATGATATTAATGCAATATCAGATTAATTAATTAATTAATTAATTAATTAATTAAT 1710

Db 1495 AGTGTATATGAATTAATGAATGCAATATCAGATTAATTAATTAATTAATTAATTAATTAAT 1554

Qy 1711 GGTAAACATCTTGATACAACTCTAAGTAAATGAGAGACCTGGTCATACAGAGGAAAC 1770

Db 1555 GGTAAACAGTCTTGATACAACTCTAAGTAAATGAGAGACCTGGTCATACAGAGGAAAC 1614

Qy 1771 TTGGTTATTTACAAAGTCAAGGGCGTTTAAATTAATTAATGTAAGTCAATCCATTTCTACA 1830

Db 1615 TTGGTTATTTACAAAGTCAAGGGCGTTTAAAGATTAATGATTAATGATTAATTTCTACA 1674

Qy 1831 CAATCTTATTTTCATTAGACTTCGATATGCTCAAAATGGTGTGAAATATCTTCCATAT 1890

Db 1675 CAATCTTATTTTCATTAGACTTCGATATGCTCAAAATGGTGTGAAATATCTTCCATAT 1734

Qy 1891 ATATCTCTTACAAATACAGAGATTAAGAAATACCACTCAACGACTCAACAACTTTT 1950

Db 1735 ATATCTCTTACAAATACAGAGATTAAGAAATACCACTCAACGACTCAACAACTTTT 1794

Qy 1951 TCTGGTACAAATTAATTAATTTTACATACGGAATTTTGGGTATTTCCATTTTCCAGT 2010

Db 1795 TCTGGTACAAATTAATTAATTTTACATACGGAATTTTGGGTATTTTCCATTTTCCAGT 1854

Qy 2011 ACAGTACATTAACCTTTAAATCGAAACATACATTTTATTTAATTCGTCAGATGATATCA 2070

Db 1855 ACAGTACATTAACCTTTAAATCGAAACATACATTTTATTTAATTCGTCAGATGATATCA 1914

Qy 2071 AATTCAAATTTTAAATCATTGATTAATTAATTTAATTAATTAATTAATTAATTAATTAAT 2130

Db 1915 AATTCAAATTTTAAATCATTGATTAATTAATTTAATTAATTAATTAATTAATTAATTAAT 1974

Qy 2131 AATGAGAAAAACAAAAATTTGAACATCTCCAAACAAAAATTAATTAATTTTTCACAAAT 2190

Db 1975 AATGAGAAAAACAAAAATTTGAACATCTCCAAACAAAAATTAATTAATTTTTCACAAAT 2034

Qy 2191 CATACAAAAACACTTGA 2208

Db 2035 CATACAAAAATTAATTTA 2052

RESULT 4

US-09-756-526A-3

Sequence 3, Application US/09756526A

Patent No. US20020038005A1

GENERAL INFORMATION:

APPLICANT: Jana, Wojciechowska

APPLICANT: Evgeny, Lewitin

APPLICANT: Ludmila, Revina

APPLICANT: Igor, Zalunin

APPLICANT: Galina, Cheestukhina

TITLE OF INVENTION: NOVEL DELTA-ENDOTOXINS AND NUCLEIC ACID SEQUENCES CODING THE REFOR

FILE REFERENCE: S-30913A

CURRENT APPLICATION NUMBER: US/09/756,526A

PRIOR FILING DATE: 2001-01-08

PRIOR APPLICATION NUMBER: US 60/175,158

PRIOR FILING DATE: 2000-01-07

NUMBER OF SEQ ID NOS: 5

SOFTWARE: PatentIn version 3.0

SEQ ID NO 3

LENGTH: 4896

TYPE: DNA

ORGANISM: Bacillus thuringiensis

**FEATURE:**

**NAME/KEY:** source

**LOCATION:** (1)..(4896)

**OTHER INFORMATION:** *Bacillus thuringiensis supsp. finitimus strain VKPM B-1161 (taxon*

**OTHER INFORMATION:** 29337

**NAME/KEY:** CDS

**LOCATION:** (1129)..(4458)

**OTHER INFORMATION:** product: Cry28Aa1 delta-endotoxin

**PUBLICATION INFORMATION:**

**AUTHORS:** Wojciechowska, et al.

**TITLE:** Two novel delta-endotoxin gene families cry26 and cry28 from *Bacillus thuringiensis* sp. *finitimus*

**JOURNAL:** FEBS Lett.

**VOLUME:** 453

**ISSUE:** 12

**PAGES:** 46-48

**DATE:** 1999-06-18

**DATABASE ENTRY DATE:** --- -- --

US-09-756-526A-3

**Query Match**      7.4%; Score 163.8; DB 3; Length 4896;  
Best Local Similarity 55.9%; Pred. No. 1.7e-20;  
Matches 405; Conservative 0; Mismatches 272; Indels 48; Gaps 3;

**Qy**    306   AGGTCCTTAATTAATATCTTTGGTACCTTAATCATCTGTCTTTGGCCCGGGAGAACCA   365  
         |||||  
**Db**    1293   AGGTGTGTACTTATATCATCTTTGGACCTTGCTCCGTCCTTGGCCTGTATCCAGAGA   1352  
         |||||  
**Qy**    366   AGA---CAAAACAGTAGGACA CAATTATTAATAATGGAGAAATTTTGTGTATACACC   422  
         |||||  
**Db**    1353   AGATTCAAAAAAAATTTGGTCA CAATTTATGAACACGGAGAAACCTTTAAATCAAAC   1412  
         |||||  
**Qy**    423   GTTAACAGAAAGCATMAAACAGCTAAAGTTACAAACCTTTGAAGAGATTAGACAAATATY   482  
         |||||  
**Db**    1413   AATTTCTACAGCTGTAAAGAAATAGCATTAAGTCTAAATGTTTTAAAGATGTATY   1472  
         |||||  
**Qy**    483   ACAAGCTATATATACAGCTTATGATGATTGGAGAAAAATTTAAAAAGATCACAGCTCCGG   542  
         |||||  
**Db**    1473   AACGTACTATGAAAGACATTTAATGATTGGAGAGAAATCCAAGTCCAATATCTGCCAG   1532  
         |||||  
**Qy**    543   ATTACCAACCATCATCAGCATTAACAACAAGCTGCCTTGACTCTTAAATAVACGATTGTGANA   602  
         |||||  
**Db**    1533   ATTTGTATCACGA-----GATTTGAAAA   1556  
         |||||  
**Qy**    603   TGTTCACAATGATTTTATTCGAGAAATACCTGTGTTCCAACTTGAAACTTATPAAAAACGCT   662  
         |||||  
**Db**    1557   CGCTCATTTCAATTTTGTATGAGCAATATGCAACAACCTCCACCTCCAGATGACACATT   1616  
         |||||  
**Qy**    663   ATTAACCTATTTATTCGCCAAGCTGCTATATTTCAITTAATTAATTATTAACAAGGTGC   722  
         |||||  
**Db**    1617   ATTTATTAAGTTGCTATATACAGAAAGCTGCAAAATTTACATTTGAATTTATTAATCAAGGTGT   1676  
         |||||  
**Qy**    723   TGAATTTGGCTATGTAATGGAATGACAGATATATACCTCTTCAAAATGAACCTAATGCTGG   782  
         |||||  
**Db**    1677   ACAATTCGGAGTCAATGGAATGACAGATCAACACATCAACAAT-----GTGAA   1727  
         |||||  
**Qy**    783   AACATCAGATGACTATTAATACTTTAAAGAAAAATATACCTTAATATAGTAATATG   842  
         |||||  
**Db**    1728   GTCATCAGGTACTTATTTATGAGAGACTATTGTATATATTTGAAAGATATTAATTAATG   1787  
         |||||  
**Qy**    843   TGCAAAATACCTATAGAACAGACTAAAAAATCTTGAAGACGAACCAAAATATGAATGAG   902  
         |||||  
**Db**    1788   CACCAAGACATPACATTAAGAGATTGAAATCACTTTAAAGAAATTCAGAAAAAATCAATGGGA   1847  
         |||||  
**Qy**    903   TATATTTATAGACTATTCGAGATATATGACACTTACTGTATTTGATACACTCTCATTT   962  
         |||||  
**Db**    1848   TGCTATTAACACATATCGTCGAAATGACCTTAATTTGATTTGATCTTGTGCGCACTTT   1907  
         |||||  
**Qy**    963   TTCTTTATATGATATATAAAAGATATAGAGATTCAATAGAGAAATAGAAATTAAGAGCAT   1022  
         |||||  
**Db**    1908   TCCTTTTATATATATACGTGCTTTCCAGAGAGATAGAACTGAATTAACAAGAGGT   1967  
         |||||  
**Qy**    1023   TAAGA 1027

Db 1968 TTATA 1972

RESULT 5  
US-10-345-020-3  
Sequence 3, Application US/10345020  
Publication No. US20030150018A1  
GENERAL INFORMATION:  
APPLICANT: Jana, Wojciechowska  
APPLICANT: Evgeny, Lewitin  
APPLICANT: Ludmila, Revina  
APPLICANT: Igor, Zalunin  
APPLICANT: Galina, Chestukhina  
TITLE OF INVENTION: NOVEL DELTA-ENDOTOXINS AND NUCLEIC ACID SEQUENCES CODING THEREFOR  
FILE REFERENCE: S-30913B  
CURRENT APPLICATION NUMBER: US/10/345,020  
CURRENT FILING DATE: 2003-01-15  
PRIOR APPLICATION NUMBER: US 60/175,158  
PRIOR FILING DATE: 2000-01-07  
NUMBER OF SEQ ID NOS: 5  
SOFTWARE: PatentIn version 3.0  
SEQ ID NO 3  
LENGTH: 4896  
TYPE: DNA  
ORGANISM: *Bacillus thuringiensis*  
FEATURE:  
NAME/KEY: source  
LOCATION: (1)..(4896)  
OTHER INFORMATION: *Bacillus thuringiensis* supsp. *finitimus* strain VKPM B-1161 (taxon  
OTHER INFORMATION: :29337  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (1129)..(4458)  
OTHER INFORMATION: product: Cry28a1 delta-endotoxin  
PUBLICATION INFORMATION:  
AUTHORS: Wojciechowska, et al.  
TITLE: Two novel delta-endotoxin gene families cry26 and cry28 from *Bacillus thuringiensis* ssp. *finitimus*  
JOURNAL: *FEBS Lett.*  
VOLUME: 453  
ISSUE: 12  
PAGES: 46-48  
DATE: 1999-06-18  
US-10-345-020-3

Query Match 7.4%; Score 163.8; DB 6; Length 4896;  
Best Local Similarity 55.9%; Pred. No. 1.7e-20;  
Matches 405; Conservative 0; Mismatches 272; Indels 48; Gaps 3;

QY 306 AGGTGCTAATAATATCTTTGTACCTATATCACTGCTTTTGGCCCGGGAGAAC 365  
DB 1293 AGGTGTGTAATTATATCTTTGTACCTATATCACTGCTTTTGGCCCGGGAGAAC 1352  
QY 366 AGA---CAAAACGATGACACATTTATTAATGGGAAATTTTGTGATACAC 422  
DB 1353 AGATCCAAAAAATTTGTGTACATTTATGAAACGAGAGACCTTTTAATCAAC 1412  
QY 423 GTTAACAGAAAGCATAAACAGCTAAAGTTACAACTTTAGAGATTAGACAAATAT 482  
DB 1413 AATTCTACAGCTGTAAAAAGAAATGCACTTACATCTAAATGTTTAAAGATGAT 1472  
QY 483 ACAAGCATATATACAGCATTTAGATGATGGAGAAATTTAAAGACTACAACTCTGG 542  
DB 1473 AACGTAATGAAAGCATTTAATGATGGAGAAATCAAGTGCAAAATACCTGCAAG 1532  
QY 543 ATTACCAATCATATAGCATTAACAACAAGCTGCTTGACTCTTAATAATGATTTGAGA 602  
DB 1533 ATTGTATCACAGA-----GATTTGAAA 1556  
QY 603 TGTTCACATGATTTTATTCGAGAAATACCTGTTTCCACTTGAACTTAAATGAGCT 662  
DB 1557 CGCTCATTTCAATTTTGTAGCAATATGCAACACTCCAACTTCCGACGTATGACACT 1616

QY 663 ATTAATCTAATTATATGCGAAGCTGCTAAATTTTCATTTAAATTTATTAACAAGGTGC 722  
DB 1617 ATTAATGAGTTGCTATACAGAGCTGCAGAAATTTACATTTGATTAATCAAGGTGT 1676  
QY 723 TGAATGGCTGATGAATGGAATGCAATATACATCCCTTCCAAATTAACCTAATGCTGG 782  
DB 1677 ACAATTCGCGATCAATGAAGATGCAATCAACCATTCACCAAT-----GTTGAA 1727  
QY 783 AACATCAGATGACTATTATTAATCTTTAAAGAAATATACCTTAATATAGTACTATG 842  
DB 1728 GTCATAGGTAATCTATATATGACAGCATTTGATATATTTGAAAGATATATTAATTG 1787  
QY 843 TGCATAATCTATTAAGAACAGACTTAAAAATCTTTAGAGCAACCAATATGAATGAG 902  
DB 1788 CACCAAGACATACATTAAGGATGGAATCACTTAAGAAATCAGAAAAATCACATGGGA 1847  
QY 903 TATATTTAATGACATATGGAATATATGACATTAATCTGTTATATACATCTTCAT 962  
DB 1848 TGCTTAATACATATATGCTGAGAAATGACCTTAATGTTGATCTTGTCGCACTTT 1907  
QY 963 TTCTTATATGATTAATAAAGATATAGAGATTCAATAGAGAAATAGAGTAAAGCAT 1022  
DB 1908 TCCTTTTATGATATATAGTGTGTTTCCAGAGAGTGAACATTAACAAGAGGT 1967  
QY 1023 TAAGA 1027  
DB 1968 TTATA 1972

RESULT 6  
US-10-342-821-3  
Sequence 3, Application US/10342821  
Publication No. US20030154510A1  
GENERAL INFORMATION:  
APPLICANT: Jana, Wojciechowska  
APPLICANT: Evgeny, Lewitin  
APPLICANT: Ludmila, Revina  
APPLICANT: Igor, Zalunin  
APPLICANT: Galina, Chestukhina  
TITLE OF INVENTION: NOVEL DELTA-ENDOTOXINS AND NUCLEIC ACID SEQUENCES CODING THEREFOR  
FILE REFERENCE: S-30913C  
CURRENT APPLICATION NUMBER: US/10/342,821  
CURRENT FILING DATE: 2003-01-15  
PRIOR APPLICATION NUMBER: US 60/175,158  
PRIOR FILING DATE: 2000-01-07  
NUMBER OF SEQ ID NOS: 5  
SOFTWARE: PatentIn version 3.0  
SEQ ID NO 3  
LENGTH: 4896  
TYPE: DNA  
ORGANISM: *Bacillus thuringiensis*  
FEATURE:  
NAME/KEY: source  
LOCATION: (1)..(4896)  
OTHER INFORMATION: product: Cry28a1 delta-endotoxin  
PUBLICATION INFORMATION:  
AUTHORS: Wojciechowska, et al.  
TITLE: Two novel delta-endotoxin gene families cry26 and cry28 from *Bacillus thuringiensis* ssp. *finitimus*  
JOURNAL: *FEBS Lett.*  
VOLUME: 453  
ISSUE: 12  
PAGES: 46-48  
DATE: 1999-06-18  
US-10-342-821-3

Query Match 7.4%; Score 163.8; DB 6; Length 4896;

Best Local Similarity 55.9%; Pred. No. 1.7e-20;  
Matches 405; Conservative 0; Mismatches 272; Indels 48; Gaps 3;

QY	306	AGGTCCTTAATTAATCTTTTGGTACCTTAATCACTGCTTTTGACCCGCGGGGAACA	365
Db	1293	AGGTGTGATCTTAATCATCTTTGGAACTTGCTGCCCTCTTTGGCTGATCCAGAGA	1352
QY	366	AGA---CAAAACAGTAGGACA CAATTATTAATAATGGAGAAATTTTGTGTATCAC	422
Db	1353	AGATCCAAAAAAATTTGGTGC CAATTTATGAA CACGGAGAAACCTTTTAATCAAC	1412
QY	423	GTTAACAGAAAGCATTAACAGCTAAAGTTACAACTTTAGAGAGTTTACAAATTT	482
Db	1413	AAATTCACACTGTAAAAAAGAAATGCAATGCTCATCTAAATGTTTAAAAAGATGAT	1472
QY	483	ACAAAGCTATATATACAGATTAGATGATGGAGAAATTTAAAGAACTTCAAGCTCCG	542
Db	1473	AACGACTATGAAAGACATTTTAATGATTTGGAGAAATCCAACTGCAATATCGCAG	1532
QY	543	ATTACCAACCATCATCAGATTACA CAAGCTGCTTGA CTTTAAATATGATTTGAGAA	602
Db	1533	ATTGGTATCAGA-----GATTTGAAA 1556	
QY	603	TGTTCACATGATTTTATTCAGAAATACCTGCTTCCAACTTGAACTTATTAAGCT	662
Db	1557	CGCTCATTTCAATTTGTAGCAATATGCCAACCTCCAACTCCACGATATGACACAT	1616
QY	663	ATTACTACCTATTTATGGCAAGCTGTAAATTTTCATTTAAATTAATTATTAACAAGG	722
Db	1617	ATTATTAATGTTGCTATACAGAAAGCTGC AAATTTACATTTGAATTAATTAACAAAGGT	1676
QY	723	TGAATGGCTATGATGAAATGCAATGCAATATACCTCCACAAATTTGAACCTAATGCTG	782
Db	1677	ACAATGGCGATGCATGGAATGCGATCAACACATCACCAAT-----GTGAA 1727	
QY	783	AACATCAGATGCTATTTAAACCTTTAAAAAGAAATATACCTTAATATGTAATCTAT	842
Db	1728	GTCAATCAGGTACTTAATATGACGAGCTATGGTAAATATGAAAAAGTATATTAATTT	1787
QY	843	TGCAATATCCATATGAA CAGACCTAAAAATCTTAGAGACGACCAATATGAATGAG	902
Db	1788	CACCAAGACATCCATAAAGATGGAATCACCTTAAAGAAATCAGAAAAAATCACATGG	1847
QY	903	TATATTTATGACTATCGAAGATATATGACCATTACTGTATTAAGTATCACTCTCA	962
Db	1848	TGCTTATACATATGCTCGAAGAAATACCTTAATTTATTTGATCTTGTGCGCACTT	1907
QY	963	TTCTTTATGATATTA AAAAGATATAGAGATTCATAGAGAAATGAAAGTAAAGCAT	1022
Db	1908	TCTTTTATATATACGTGCTTTCCAAAGGAGATGAACCTAAGATTAACAAAGAGGT	1967
QY	1023	TAAGA 1027	
Db	1968	TTATA 1972	

RESULT 7  
US-10-089-678-2

```

Sequence 2, Application US/10089678
Publication No. US20030017967A1
GENERAL INFORMATION:
APPLICANT: ASANO, Shinichiro
TITLE OF INVENTION: PROTEIN HAVING INSECTICIDAL ACTIVITY, DNA ENCODING THE PROTEIN, AND
TITLE OF INVENTION: NOXIOUS ORGANISM-CONTROLLING AGENT AND METHOD
FILE REFERENCE: 068821
CURRENT APPLICATION NUMBER: US/10/089,678
CURRENT FILING DATE: 2002-05-02
PRIOR APPLICATION NUMBER: JP 2000-236140
PRIOR FILING DATE: 2000-08-03
PRIOR APPLICATION NUMBER: PCT/JP01/06660
PRIOR FILING DATE: 2001-08-02
NUMBER OF SEQ ID NOS: 3
SOFTWARE: PatentIn version 3.1

```

```

; SEQ ID NO 2
; LENGTH: 3504
; TYPE: DNA
; ORGANISM: Bacillus thuringiensis
; FEATURE:
; NAME/KEY: exon
; LOCATION: (1)..(3501)
; OTHER INFORMATION:
US-10-089-678-2

```

Query Match	4.5%;	Score 98.4;	DB 5;	Length 3504;
Best Local Similarity	43.4%;	Pred. No. 3.1e-08;		
Matches 602;	Conservative 0;	Mismatches 751;	Indels 33;	Gaps 2;

QY	612	TGATTTTATTCGAGAAATATACCTGGTTTCAACTTGAACCTTATTAACGCTATTATAC	671
Dp	531	TGATTTTGTGCTAAATATCCCATCTTTTGCATATCTGCACAGAAATACATTTATATC	590
QY	672	TATTTATGCGCAGCTGCTAATTTTCACTTTAAATTTATACACAAGTGCTGAATTGGC	731
Dp	591	AGCTATATGACACAGACAGAAATTTACATTTGCTATTATTCAGAAATGCTTCCATTTTGG	650
QY	732	TGATGAATGGAATGACGATATATACATCCTTCACAATTTGAACCTAATGCTGGAACATCAGA	791
Dp	651	AGCAGAGTGC-----GGATTCCACCAGAGAAATTTCC	683
QY	792	TGACTATTATTAACCTTTTAAAGAAATATATACCTAAATATATAGTAATAATTTGCAAAATAC	851
Dp	684	CACATTTTATGATGCTCAGGAGACAGTACCGCCCAATACCTCGGATTATTTGTATAAAGTG	743
QY	852	CTATAGAACAGGACTTAAAAATCTTAAGACAGAACCAATATGAAATGGAATATTTTAA	911
Dp	744	GTATTAACACCTGGCTTAAATATTTAAAGGATACGAATGCTCGAAGTTGGCTGAAGATATCA	803
QY	912	TGACTATCGAAGATATATGACCATTTACTGTATTTAGATACCATCTCTCAATTTTCTTTATA	971
Dp	804	CCATTTCCGAAGAAATGACATTACTGATTTTATGATTTAGTACCGTTATTTCCAAACTA	863
QY	972	TGATATTAATAAGATATATAGATTCAATATGAGAGAAATGAAAGTAAAGCATTTAAGATGA	1031
Dp	864	TGACACACGTATCGATCCAAATCGAAACACGCGCCCACTTACACGGGAAGTGTATACAGA	923
QY	1032	ACTCACAAGGAATTTATACAACTGAAATATTTTGAATGCTTCTCTCACTTAAGAT	1091
Dp	924	TCCAAATGTTATTTAACAGAAACAAAGTGTGATTTGTAGGGTGTGCATTTACAG	983
QY	1092	TCAACCAATCTAGCTACGATGGAATATATTTTACACGTGCAAGTTTAAATATATTTTC	1151
Dp	984	TGATATTTCTTTTACAGAAATCGAAACCGCTGTATTCGTTCCACCAACCATTTATGATAT	1043
QY	1152	ATTTTTAGAACAAATTTATTTTATACAGAAATAACAATTTGCGGAATGCTTTAGTTGG	1211
Dp	1044	ACTCAGTGAATATGAAATTTTATTCACACMAAGCGGGGCTTCCCTTGATATATACGGAATA	1103
QY	1212	TATTTCAATCGATGACACCTACT-----TATAGCAATCTTTAACTGAAACTTTATA	1265
Dp	1104	CTTTGATATTTGGGTAGGACATTTCTATTAATAATATTAATAATATGCAATGCTCATGACATT	1153
QY	1266	TGAGAAAGAACAGGTTCAACCCACACAAATAACATATAGACCAATTTGAATCTTTTAAAGT	1325
Dp	1164	AGAAACGTAAATACGGTACGAATTACTTTAACAAATAACAGATATATGATTTAGCAAAATTA	1223
QY	1326	TTCAATTTGTAACGTATGACATACACTCTGTTTCCCTATTTCAACACACCTTTATAT	1385
Dp	1224	GGATATCTTTCAGGTTTGATCAATTAAGGGGCGGAATTAAGCTTAATTTACACGACAGGTATA	1283
QY	1386	TAAATCAAAATGAACCTTATTTTAAATGCGCTCATCTTAACACACACTCAAAATATTCAGACAG	1445
Dp	1284	TGAGATTCGGTACGCTATGTTTACACCTGCTTGACAAAGATATACAGGATCAAGATCAGTTGG	1343
QY	1446	AGGCTCTTTATCTATATATCAAAACACACACTTTTTCATTTCTGTAGAAAAAAGACTG	1505
Dp	1344	AGGTTTATACGTACTCAAAACCAACATACATATGACAGATATGTACACAAATATTCATATAC	1403

OY 1506 CAATCTAGTATGATCCAGGTGTTCCAAACTTATACTATAGTATATTTATC 1565  
DB 1404 GATTGATGAATCCCTCCAGGAATGAGCCTTAGTAGAGGATATACCATTAATTC 1463  
OY 1566 CCATTTTTCATTTATTTACTTATTTCTTATGTATGTATGATTAACAGCTACAAATTTAGATAC 1625  
DB 1464 TCATATCACTCTTATTTCTTTTCTTAAAGATGCTAGTCTCTGTAGATATGCAATCT 1523  
OY 1626 AGGTATATTAAGATGAGACACAGTAGTGTATGATATATATGCAATATGCAATTAAT 1685  
DB 1524 CCGTATATTTGCTTGAACATCGAGTGGGATGTTCAAATACAGTTTATTCAGATTA 1583  
OY 1686 AATTACATGATATCCAGCAATCAAGGTACCAATCTTGAATCAAACTGTAAGTATGA 1745  
DB 1584 AATTACTCAGATACCAAGTTGTAAGGACATATCTTAGTTCAGTACTACTGTTATTA 1643  
OY 1746 AGGACCTGTATACAGAGAGAACTTGTATTTATTAAGATCAAGGCGTTTGAAT 1805  
DB 1644 AGGTCTGTGATTTACAGAGGCAATATCTTAAAGAAACAGTATGTCGTTAGCTTA 1703  
OY 1806 TACATGTAACTCTTAATTTCTACACATCTTATTTCACTTAGACTTGTATATGCTACAA 1865  
DB 1704 TACTAGTCTCTGTAAATACCATTTATCAAAAGATATGTCGACAGAAATACGTTATGC 1763  
OY 1866 TGGTGTGGAATATCTTCTTAATATATCTTTACAAATACAGAGATATAGAAATCC 1925  
DB 1764 TTCTACTACTAATCAAGACTTTTGTAAACATTTCTGAACTGCAATTTACTATATA 1823  
OY 1926 ACCTCACAGACTCAACAACACTTTTCTGTACAAATTTATTAATTTTACAAATGAGA 1985  
DB 1824 TGTTAATTAACATGAATTAAGGAGATGATTAACATTTATATGATTTGACTTAGCAAC 1883  
OY 1986 TTTTGG 1991  
DB 1884 TATTGG 1889

RESULT 8  
US-10-089-678-3  
; Sequence 3, Application US/10089678  
; Publication No. US20030017967A1  
; GENERAL INFORMATION:  
; APPLICANT: ASANO, Shinichiro  
; TITLE OF INVENTION: PROTEIN HAVING INSECTICIDAL ACTIVITY, DNA ENCODING THE PROTEIN, A  
; FILE REFERENCE: 068821  
; CURRENT APPLICATION NUMBER: US/10/089,678  
; PRIOR FILING DATE: 2002-05-02  
; PRIOR APPLICATION NUMBER: JP 2000-236140  
; PRIOR FILING DATE: 2000-08-03  
; PRIOR APPLICATION NUMBER: PCT/JP01/06660  
; NUMBER OF SEQ ID NOS: 3  
; SOFTWARE: Patentin version 3.1  
; SEQ ID NO 3  
; LENGTH: 3690  
; TYPE: DNA  
; ORGANISM: *Bacillus thuringiensis*  
US-10-089-678-3

Query Match 4.5%; Score 98.4; DB 5; Length 3690;  
Best Local Similarity 43.4%; Pred.No. 3.1e-08;  
Matches 602; Conservative 0; Mismatches 751; Indels 33; Gaps 2;

OY 612 TGATTTTATTTGAGAAATACCTGTTTCCAACTGAACTTATAAAGCTATTAATCC 671  
DB 717 TGATTTTGTCTTAATCCATCTTTTGCATATCTGACAGAAAGTACATTAATTC 776  
OY 672 TATTATGCGAAGCTGTCTAATTTTCAATTAATTTATTAACAAGTGTGAATGGC 731  
DB 777 AGTGTATGCAAGCAGCAATTTTCAATTTGCTATTTATGAGATGCTTCAATTTTGG 836

OY 732 TGATGAATGGAATGACATATACATCTTCAAAATTTGAACCTAATGCTGGAATCAGA 791  
DB 837 AGCAGAGTGG-----GGATTTCAACACAGAGAAATTTTC 869  
OY 792 TGACTATTTATTAATCTTTTAAAGAAATATACCTAATATATAGTAACTATTTGTCAATAC 851  
DB 870 CACATTTTATGATGTCAGGTGACAGGTACCGCCCAATACCTGGAATTAATGTGTAAGTGC 929  
OY 852 CTATAGAACAGGAACTTAAATAATCTTAGAGACCAACCAATATGAATGAGATATATTTAA 911  
DB 930 GTATTAACCTGCTTAATTAATTTTAAAGGATGCAATGCTGCAAGTTGGCTGAAGATCA 989  
OY 912 TGACTATGGAATATATGACATTAATGATTAATGATATCAATCTTCAATTTTCTTATA 971  
DB 990 CCAATTTCCGAAGAAATGACATTAATGATTAATGATTAATGATTAATTTTCCAACTA 1049  
OY 972 TGATATTAATAAGATATGAGATTTCAATGAGAGAAATGAACTTAAGCAATTAAGATGA 1031  
DB 1050 TGACACACGATGATCAATCGAAACACCGCCCACTTACAGGGAAGGTATACAGA 1109  
OY 1032 ACTCACAAGAAATTTATACACTGAAATTAATTTTGTATGCTCTTCACTTAGAGT 1091  
DB 1110 TCCATATGATTTTAAACAGAGAAACAGTGTGATTTTGTAGGCTGTGTCATTAACG 1169  
OY 1092 TCAACCCATCTAGCTACGATGGAATTAATTTAAACAGTGAAGTTTAATTAATTTTC 1151  
DB 1170 TGATATTTCTTTTCAAGATCGAAAGCGCTGTATTCGTTCAACACACTATTTGATAT 1229  
OY 1152 ATTTTGAACAATTTATTTTATACGAAATAATCAAAATTTGGGAAATGCTTAGTTGG 1211  
DB 1230 ACTCAGTGAATTAATAATTTTATACAAAGACGGGCTTCCCTTGAATTAATACGATA 1289  
OY 1212 TATTTCAATGCTATGACACTACT-----TATAGCAATCTAATCTGAACTTATA 1265  
DB 1290 CTTGAATATTTGGTATGAGCACTTTATTAATTAATTAATTAATGATGCTCATACGAT 1349  
OY 1266 TGAAGAAAGAACAGTTCACCCACAAACAAACATTAAGCAATTTGAATCTTAAGT 1325  
DB 1350 AGAAGTAAATTAACGGTACGATTAATCTTAAACAAATCAAGATTAATTAAGCAAAATTA 1409  
OY 1326 TTCAATTTGATGATGACATTCACCTCTGTTTCCCTATTTCAACACACTTATAT 1385  
DB 1410 GGATATCTTCAAGGTGATCATTAATGAGGGCGGATTAAGCTAATTAACGACGATTA 1469  
OY 1386 TAATCAATTTGAATTTATTTAATGCTCATCTAACAACACTCAAAATATTCAGCAGG 1445  
DB 1470 TGAAGTTCGGATGCTGTTTACACGCTTGACAAAGATTAACAGATCAAGATCAAGTTGG 1529  
OY 1446 AGGTCTTTATCTAATTAATCAAAACACAACTTTTTCATTTCTGAGAAAAAAGACTG 1505  
DB 1530 AGGTTTTACGTAACCAAAACCAATCACTATGCAAGTATGTACAAAAATTTCAATAC 1589  
OY 1506 CAATCTAGTATGATCCAGGTGTTTCCAACTTTTAATTAATCTATATGCTATATTTATC 1565  
DB 1590 GATTGATGAATCCCTCCAGGAATGAGCCATTAGTAGAGGTATAGCCATTAATATTC 1649  
OY 1566 CCATTTTCAATTTTACTTATTTCTATGATTTGATTAACAGCTACAAATATTTAGATAC 1625  
DB 1550 TCATATACCTCTTATTTCTTTTCTAAGAAATGCTAGTAGTCTGCTAGATATGCAATCT 1709  
OY 1626 AGGTATATTAAGATGACACACAGTAGTGTATGATATATATGCAATATTCAGATTAAT 1685  
DB 1710 CCGTATTTGCTTGAACACATCGAGTGGGATGTATACAAATCAAGTTTATTCAGATTA 1769  
OY 1686 AATTACATGATATCCAGCAATCAAGGTATCAATCTTGAATCAAACTTAAAGTATTTGA 1745  
DB 1770 AATTACTCAGATACCAATTTGTAAGGACATATCTTAAGTTTCAGATCTACTGTTATTA 1829  
OY 1746 AGGACCTGTCTATCAAGAGAACTTGTATTTTCAAAAGTCAAGGCGTTTGAAGAT 1805  
DB 1830 AGGTCTGTGATTTACAGAGGCAATATCTTAAAGAAACAGATGATGCTGCTTAGCTTA 1889  
OY 1806 TACATGTGAACCTCTAATTTCTACCAATCTTATTTCACTTAGACTTGTATATGCTACAA 1865

Db 1890 TACTAGTCTCTGTAAATACACATTTACAAAGATATCTGACAGATACGTTATGC 1949  
Qy 1866 TGGTCTGGAATACTCTTCCATATATATCTTCAATACAGAGTAATGAAATCC 1925  
Db 1950 TTCTACTACTACTTACACACTTTTGTGAACAATTTCTGACCTGCATTTACTCTATTA 2009  
Qy 1926 ACCTCAACGACTCAACAACTTTTCTGTCAAAATTAATTAATTTACATACGAGA 1985  
Db 2010 TGTTAATTAACCAAGAAATTAAGGGAGATTTAACTTTAATTAACATTTGACTAGCAAC 2069  
Qy 1986 TTTTGG 1991  
Db 2070 TATTTGG 2075

RESULT 9  
US-10-312-841-1/c  
; Sequence 1, Application US/10312841  
; Publication No. US20030186277A1  
; GENERAL INFORMATION:  
; APPLICANT: Bpigenomics AG  
; TITLE OF INVENTION: Diagnose von bedeutenden genetischen Parametern innerhalb des MHC  
; FILE REFERENCE: E01/1208/WO  
; CURRENT APPLICATION NUMBER: US/10/312,841  
; NUMBER OF SEQ ID NOS: 2  
; SEQ ID NO 1  
; LENGTH: 3673778  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)  
; NAME/KEY: unsure  
; LOCATION: (3294164)  
US-10-312-841-1

Query Match 4.4%; Score 97; DB 6; Length 3673778;  
Best Local Similarity 43.1%; Pred. No. 1.1e-06;  
Matches 629; Conservative 0; Mismatches 825; Indels 7; Gaps 3;

Qy 749 ATATACATCTTCAACAAATGAACTTAATGCTGAACATCAGATGACTATTAACTTT 808  
Db 1715403 AAATATATATATTAATTAATACATTAATTAATATATATTAATTAATTAATTAAT 1715344  
Qy 809 TAAAGAAATATATCTTAATATAGTAACATTTGTGCAATACCTATAGAACAGACTAA 868  
Db 1715343 ATATATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1715284  
Qy 869 AAATCTTAGAGACCAACCAATNGAATGAGATATTTTAATGACATGCAAGATATA 928  
Db 1715283 AAATATATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1715224  
Qy 929 TGACCATTTACTGTATTTAGATACCATCTCAATTTTCTTTATATGATTTAAAGATATA 988  
Db 1715223 TATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1715164  
Qy 989 GAGATTCATAGAGGAATAGAGTAAGTAAGGCAATTAAGATGAACCTCAAGAGAAATTT 1048  
Db 1715163 TAAATATATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1715104  
Qy 1049 ATACAACTGAATTAATTTTGAATGCTCTTCCCAACTTGAAGTCAACCAATCTAGCTA 1108  
Db 1715103 ATATATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1715044  
Qy 1109 CGATGGAATATTAATTTAAGCTGCAAGTTTAAATTTTCAATTTTGAACAATTTA 1168  
Db 1715043 ATATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1714984  
Qy 1169 TTTTATATACGAAATATCAAAATTTGGGAATCGTTAGTGGTATTTCTAATCGTGATG 1228  
Db 1714983 ATATATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1714927

Qy 1229 CACCTACTTAATAGCAATACTATTAACCTGAACCTTATATGAGAAAGAACAGGTCACCA 1288  
Db 1714926 TATATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1714867  
Qy 1289 CAACAAAACATTAAGACACTTTGATCTTATA--AGTTTCAATTTGTAAGTATAGC 1345  
Db 1714866 AAATATATACATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1714807  
Qy 1346 AATCACTCCCTGTTCCCTATTCACACACTTATTAATTAATTAATTAATTAATTAAT 1405  
Db 1714806 AATATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1714747  
Qy 1406 TAAATGCTCATCTTAACAACACACTCAATATTTACAGAGAGGCTTATCTAATTTATC 1465  
Db 1714746 TAAATATATACATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1714687  
Qy 1466 AAAACACAACTTTTTCATTTCTGAGAAAAAAGACTGCAATCTAGTTATGATCCAG 1525  
Db 1714686 ACATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1714627  
Qy 1526 GTTGTCACCAACTTTTAATTAATTAATTAATTAATTTTATCCATTTTCAATTTACTT 1585  
Db 1714626 AATATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1714567  
Qy 1586 ATTCCTATGATTTGATTTACAGCTACCAATATTTAGATACAGGTGATTTAGATGACAC 1645  
Db 1714566 TATACATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1714507  
Qy 1646 ACAGTAGTGTGAATAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1705  
Db 1714506 ATATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1714447  
Qy 1706 TCAAAAGTAACTTTGATTAACAACTTAAGGTAATTTGAAGACCTGCTATACAGAG 1765  
Db 1714446 TATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1714387  
Qy 1766 GAACTTGTTTATTTTACAAAGTCAAGGCGCTTGAATTAACATGTAAGTCACTCTTAAT 1825  
Db 1714386 ATATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1714327  
Qy 1826 CTACCAACTCTTATTTCTTACCTTGAAGTCTGATTAATGCTCAAAATGCTGGAATCTTTC 1885  
Db 1714326 TATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1714267  
Qy 1886 CTATATATCTCTTACAAATACAGAGTAATTAAGAAATCCACTCAAGACTCAACACA 1945  
Db 1714266 ATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1714207  
Qy 1946 CTTTTCGTGACAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2005  
Db 1714206 ATATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1714148  
Qy 2006 CAACTACGTAACATTAACCTTAAATCGAACAATACATTAATTAATTAATTAATTAAT 2065  
Db 1714147 ATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1714088  
Qy 2066 TATCAATTTCAATTTTATCTGATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2125  
Db 1714087 TATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1714028  
Qy 2126 ACCAAATTAAGGAAGAAACAAATTTAGAACTATCCAAACAAATTAATTAATTAATTAAT 2185  
Db 1714027 ATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1713968  
Qy 2186 CAAATCATACAAAACACTTT 2206  
Db 1713967 TAAATATATTAATTAATTAAT 1713947

RESULT 10  
US-10-311-455-2128/c  
; Sequence 2128, Application US/10311455









Db 5762 TTTTATATACAAAATTAAACAATTAAATTTTCAATATATTTATTCACATCATT 5703  
Qy 2030 ATCGAAACATACCATTTATTTATTCGTGAGATGATCAATTCATTTATTCATG 2089  
Db 5702 ATCATCAATTTTCATTCATTTTATTTTATTTTATTTTATTTTATTTTATTTT 5643  
Qy 2090 ATAAATTTGAATTTATACCAATTCCTTCCTGATGACCAAAATATGAAAAACAAAAT 2149  
Db 5642 AAAATTTTACAAAAAATTTTCAATTTTCAATTTATTTTAAATTTTCAATTTTC 5583  
Qy 2150 TAGAAATCTCCAAACAAAATTTTATACATTTTTCACAAATCATTCACAAAACAC 2203  
Db 5582 AAAATTTTATTTATTCATTTATTAATTTTATTAATTTTAAATTTTAAATTTTAAAC 5529

## RESULT 12

US-09-826-660-5  
; Sequence 5, Application US/09826660  
; Patent No. US20010026940A1  
; GENERAL INFORMATION:  
; APPLICANT: Cardineau, Guy A.  
; APPLICANT: Steilman, Steven J.  
; APPLICANT: Narva, Kenneth E.  
; TITLE OF INVENTION: Plant-Optimized Genes Encoding Pesticidal Toxins  
; FILE REFERENCE: MA-714XCD1  
; CURRENT APPLICATION NUMBER: US/09/826,660  
; PRIOR FILING DATE: 2001-04-05  
; PRIOR APPLICATION NUMBER: 09/178,252  
; PRIOR FILING DATE: 1998-10-23  
; PRIOR APPLICATION NUMBER: 60/065,215  
; PRIOR FILING DATE: 1997-11-12  
; PRIOR APPLICATION NUMBER: 60/076,445  
; PRIOR FILING DATE: 1998-03-02  
; NUMBER OF SEQ ID NOS: 27  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 5  
; LENGTH: 3522  
; TYPE: DNA  
; ORGANISM: *Bacillus thuringiensis*  
US-09-826-660-5

Query Match 3.9%; Score 85.8; DB 3; Length 3522;

Best Local Similarity 53.8%; Pred. No. 7.3e-06;  
Matches 220; Conservative 0; Mismatches 162; Indels 27; Gaps 1;

Qy 589 ATACGATTTGAGAAATGTTACAAATGATTTTATTCGAGAAATACCTGTTCCAACTGAA 648  
Db 388 ATTGATTTGCTAATACAGAGAGCGCTTAAATACAGCAATTAATTTTACACTTACA 447  
Qy 649 ACTTATAAAGCGTATTACTACTATTATTTATGCGCAAGCTGCTAATTTTCAATTTA 708  
Db 448 AGTTTGAATCCCTCTTTTATCGCTATGTTCAAGCGGCAATTTTACATTTTACTA 507  
Qy 709 TTACAAGAGGCTGATGCTGATGCTGATGATGAGATGAGATATACCTTCCAAAT 768  
Db 508 TTAAGAGCGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCT 567  
Qy 769 GAACCTAATGCTGAAATCATGATGATGATGATGATGATGATGATGATGATGAT 828  
Db 568 C-----ATTATATATGATTTAATTAATTTTAAATTTTAACTTATTCATTA 600  
Qy 829 TATAGTAACTATTGTAATTTGTAATTTGTAATTTGTAATTTGTAATTTGTAATTT 888  
Db 601 TATAGTAACTATTGTAATTTGTAATTTGTAATTTGTAATTTGTAATTTGTAATTT 660  
Qy 889 AATATGAAATGAGATATTTTATGATGATGATGATGATGATGATGATGATGATGAT 948  
Db 661 ACTGACAAATGGGCAAAATTTCAATTTCAATTTCAATTTCAATTTCAATTTCAAT 720  
Qy 949 ACCATCTCTCAATTTTCTTTATATGATATATAAAGATATAGAGATTCAA 997  
Db 721 ATGCTGCTCTTTTTCGAACTAGATGTTAGAACATATTCATTTCAA 769

## RESULT 13

US-09-837-961-7  
; Sequence 7, Application US/09837961  
; Publication No. US2004005860A1  
; GENERAL INFORMATION:  
; APPLICANT: Payne, Jewel  
; APPLICANT: Sick, August  
; TITLE OF INVENTION: No. US2004005860A1 *Bacillus thuringiensis* Isolate Active Against  
; FILE REFERENCE: MA-43CDPFD3  
; CURRENT APPLICATION NUMBER: US/09/837,961  
; PRIOR FILING DATE: 2001-04-19  
; PRIOR APPLICATION NUMBER: US 09/521,344  
; PRIOR FILING DATE: 2000-03-09  
; PRIOR APPLICATION NUMBER: US 08/933,891  
; PRIOR FILING DATE: 1997-09-19  
; PRIOR APPLICATION NUMBER: US 08/356,034  
; PRIOR FILING DATE: 1994-12-14  
; PRIOR APPLICATION NUMBER: US 08/210,110  
; PRIOR FILING DATE: 1994-03-17  
; PRIOR APPLICATION NUMBER: US 07/865,168  
; PRIOR FILING DATE: 1992-04-09  
; PRIOR APPLICATION NUMBER: US 07/451,261  
; PRIOR FILING DATE: 1989-12-14  
; PRIOR APPLICATION NUMBER: US 371,955  
; PRIOR FILING DATE: 1989-06-27  
; NUMBER OF SEQ ID NOS: 8  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 7  
; LENGTH: 3522  
; TYPE: DNA  
; ORGANISM: *Bacillus thuringiensis*  
US-09-837-961-7

Query Match 3.9%; Score 85.8; DB 3; Length 3522;

Best Local Similarity 53.8%; Pred. No. 7.3e-06;  
Matches 220; Conservative 0; Mismatches 162; Indels 27; Gaps 1;

Qy 589 ATACGATTTGAGAAATGTTACAAATGATTTTATTCGAGAAATACCTGTTCCAACTGAA 648  
Db 388 ATTGATTTGCTAATACAGAGAGCGCTTAAATACAGCAATTAATTTTACACTTACA 447  
Qy 649 ACTTATAAAGCGTATTACTACTATTATTTATGCGCAAGCTGCTAATTTTCAATTTA 708  
Db 448 AGTTTGAATCCCTCTTTTATCGCTATGTTCAAGCGGCAATTTTACATTTTACTA 507  
Qy 709 TTACAAGAGGCTGATGCTGATGCTGATGATGAGATGAGATATACCTTCCAAAT 768  
Db 508 TTAAGAGCGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCT 567  
Qy 769 GAACCTAATGCTGAAATCATGATGATGATGATGATGATGATGATGATGATGAT 828  
Db 568 C-----ATTATATATGATTTAATTAATTTTAAATTTTAACTTATTCATTA 600  
Qy 829 TATAGTAACTATTGTAATTTGTAATTTGTAATTTGTAATTTGTAATTTGTAATTT 888  
Db 601 TATAGTAACTATTGTAATTTGTAATTTGTAATTTGTAATTTGTAATTTGTAATTT 660  
Qy 889 AATATGAAATGAGATATTTTATGATGATGATGATGATGATGATGATGATGATGAT 948  
Db 661 ACTGACAAATGGGCAAAATTTCAATTTCAATTTCAATTTCAATTTCAATTTCAAT 720  
Qy 949 ACCATCTCTCAATTTTCTTTATATGATATATAAAGATATAGAGATTCAA 997  
Db 721 ATGCTGCTCTTTTTCGAACTAGATGTTAGAACATATTCATTTCAA 769

## RESULT 14

US-10-825-751-7  
; Sequence 7, Application US/10825751  
; Publication No. US20040194165A1  
; GENERAL INFORMATION:

```
; APPLICANT: Payne, Jewel
; APPLICANT: Sick, August J.
; TITLE OF INVENTION: Novel Bacillus thuringiensis Isolate Active Against Lepidopteran
; TITLE OF INVENTION: Pests, and Genes Encoding Novel Lepidopteran-Active Toxins
; FILE REFERENCE: MA-43CDF204
; CURRENT APPLICATION NUMBER: US/10/825,751
; PRIOR FILING DATE: 2004-04-16
; PRIOR APPLICATION NUMBER: US 09/837,961
; PRIOR FILING DATE: 2001-04-19
; PRIOR APPLICATION NUMBER: US 09/521,344
; PRIOR FILING DATE: 2000-03-09
; PRIOR APPLICATION NUMBER: US 08/933,891
; PRIOR FILING DATE: 1997-09-19
; PRIOR APPLICATION NUMBER: US 08/356,034
; PRIOR FILING DATE: 1994-12-14
; PRIOR APPLICATION NUMBER: US 08/210,110
; PRIOR FILING DATE: 1994-03-17
; PRIOR APPLICATION NUMBER: US 07/865,168
; PRIOR FILING DATE: 1992-04-09
; PRIOR APPLICATION NUMBER: US 07/451,261
; PRIOR FILING DATE: 1989-12-14
; PRIOR APPLICATION NUMBER: US 371,955
; PRIOR FILING DATE: 1989-06-27
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 7
; LENGTH: 3522
; TYPE: DNA
; ORGANISM: Bacillus thuringiensis
; US-10-825-751-7
```

```
Query Match 3.9%; Score 85.8; DB 8; Length 3522;
Best Local Similarity 53.8%; Pred. No. 7.3e-06;
Matches 220; Conservative 0; Mismatches 162; Indels 27; Gaps 1;
```

```
QY 589 ATACGATTGAGATGTTTCAAGATGTTTATTCGAGAAATCCTGGTTCCAACTTGA 648
DB 388 ATTCGATTGCTAATACAGACGACCTTAAATACAGCAATTAATTTTACACTTACA 447
QY 649 ACTTATTAAGAGCTTACTACTTATTTATGGCGAAGCTGCTTAATTTTCAATTAATTTA 708
DB 448 AGTTTGAATCCCTCTTTTATCGGCTATGTTCAAGGCGCAATTTTCAATTTTACACTA 507
QY 709 TTACAACAGAGTGTGAAATGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 768
DB 508 TTAAGAGACGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 567
QY 769 GAACCTAATGCTGAACATCAGATGATGATGATGATGATGATGATGATGATGATGATGAT 828
DB 568 C-----ATTATATGATTTAATTAATCTTATTCATGTA 600
QY 829 TATAGTACTATTTGCGCAATTCCTATATAGAACAGACATTAATAATCTTAGAGACAGACA 888
DB 601 TATAGGAACAATGTTTGGACATACATCAATCAAGATTTAGAAACTTAAAGAGTACTAAT 660
QY 889 AATATGAATGAGATATTTTAAATGACTATGAGATGATGATGATGATGATGATGATGATGAT 948
DB 661 ACTGACAAATGGGCAAGATTCATCATGTTTAGAGAGATTTTACCTTACTGTATTAAT 720
QY 949 ACCATCTCTCAATTTTCTTATATGATATTAATAAGATATAGAGATTCAA 997
DB 721 ATCGTGTCTCTTTTCCGAACTAGATGTTAGAACATATCCAAATTCAA 769
```

```
RESULT 15
US-10-311-455-1280/c
; Sequence 1280; Application US/10311455
; Publication No. US20030143606A1
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIERENBROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with the Immune System by D-term
```

```
; TITLE OF INVENTION: Cytosine methylation
; FILE REFERENCE: 5013.1014
; CURRENT APPLICATION NUMBER: US/10/311,455
; CURRENT FILING DATE: 2002-12-16
; PRIOR APPLICATION NUMBER: PCT/EP01/07537
; PRIOR FILING DATE: 2001-07-02
; PRIOR APPLICATION NUMBER: DE 10032529.7
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DE 10043826.1
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 2424
; SEQ ID NO 1280
; LENGTH: 6175
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
; US-10-311-455-1280
```

```
Query Match 3.7%; Score 81.2; DB 6; Length 6175;
Best Local Similarity 42.7%; Pred. No. 6.8e-05;
Matches 593; Conservative 0; Mismatches 788; Indels 9; Gaps 3;
```

```
QY 800 ATAACTTTTAAAGAAATATACCTAATATATGTAATCTATTTGCAAAATCTATAGAA 859
DB 5961 AAATTAATATCTAATATATATATATTAACAAACGTATTAACATCAATTAATATACAT 5902
QY 860 CAGACCTAAATAATCTTGAACACGACCAATATGAATATGATATTTATGACTATG 919
DB 5901 AAACAAATATATATATATATATATATATATATATATATATATATATATATATAT 5842
QY 920 GAAGATATATGACCATCTATCTATATATATATATATATATATATATATATATATAT 979
DB 5841 ATAAATTAACAAATATATATATATATATATATATATATATATATATATATATATAT 5782
QY 980 AAAGATATAGATTCATATATATATATATATATATATATATATATATATATATATATAT 1039
DB 5781 AAACATATATATATATATATATATATATATATATATATATATATATATATATATAT 5722
QY 1040 GAGAAATTTATACAGTAAATATATATATATATATATATATATATATATATATATATAT 1099
DB 5721 ATAAACATATATATATATATATATATATATATATATATATATATATATATATATAT 5662
QY 1100 ATTAGTACGATGGAATATATATATATATATATATATATATATATATATATATATATAT 1159
DB 5661 AACATATATATATATATATATATATATATATATATATATATATATATATATATATAT 5602
QY 1160 AACATTTATTTTATATACAGAAATATCAAAATTCGGGAATCGTTAGTGTATTTCTA 1219
DB 5601 AAATATATATATATATATATATATATATATATATATATATATATATATATATATAT 5542
QY 1220 ATCGTATGACCTTCTATATAGCAATCTATATACGAAATCTTATATATATATATATAT 1279
DB 5541 ACATTAATATATATATATATATATATATATATATATATATATATATATATATATAT 5482
QY 1280 GTTACACCAACAAACAAATATAGACCATTTGAATCTTATATATATATATATATATAT 1339
DB 5481 AACATATATATATATATATATATATATATATATATATATATATATATATATATATAT 5422
QY 1340 ATAGCAATCACTCTGTTTCCCTATTAACCAACACTTATATATATATATATATATATAT 1399
DB 5421 AAAATATATATATATATATATATATATATATATATATATATATATATATATATATAT 5362
QY 1400 TTTATATTAATGCTCATCTAACAACACTCAATATATATATATATATATATATATATAT 1459
DB 5361 TATATATATATATATATATATATATATATATATATATATATATATATATATATATAT 5306
QY 1460 ATATATCAAAACAACTTTTTCATTTCTAGAAATATATATATATATATATATATATATAT 1519
DB 5305 AATATATATATATATATATATATATATATATATATATATATATATATATATATATAT 5246
QY 1520 ATCCAGTTGTCACCAACTTATATATATATATATATATATATATATATATATATATATAT 1579
```

Db 5245 ATATTAATTAATATATATTAATTAACATATATATTAATTAACAAAAATATATATACAAA 5186  
Qy 1580 TTACTTATTCCTANGTGAATGGATTACAGTACAAATATATAGATACAGGTATTTAGGAT 1639  
Db 5185 ATAAACATATATAAATATAAATAAATTAACAATATATAAATATAAACAATATATTTATAC 5126  
Qy 1640 GGACACACAGTAGTGTGATGATTAATTAATGCAATATCAGATTAATTAATTAACAATGATCC 1699  
Db 5125 AAACATATATTAACAAATATTAACAAATATATTAATTAACCTAATATTAATTAATTA 5066  
Qy 1700 CAGCAATCAAAAGGTACAACTGTGTACAACTGTAAGTAATGGAAGACTGTGCATA 1759  
Db 5065 CAAATATATTAACATTAATTAATTAACCTAATATTAATTAATTAATTAACAAATTAATA 5006  
Qy 1760 CAGGAGAACTGTGTTATTTATTAACAAGTCAAGGCGTTAGAAATTTACATGTGAACTC 1819  
Db 5005 AACATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATC 4946  
Qy 1820 CTAAATTCACCAATCTTATTTTCATTAGACTTCGATATGCTACAAATGTGTGGAATA 1879  
Db 4945 TAAA--TATATATACATATATATTAATTAATTAATTAATTAATTAATTAATTAATTA 4889  
Qy 1880 CTCTTCTAATATATCTCTTACAAATACAGGAGTATAGAAATCCACCTCAAGACTCA 1939  
Db 4888 AATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 4829  
Qy 1940 ACAACACTTTTCTGTACAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTC 1999  
Db 4828 TAAATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 4769  
Qy 2000 AATTTCAGTACGTAACATTACCTTAAATCGAAACATACATTATATTAATGCTG 2059  
Db 4768 AAATATATCGTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 4709  
Qy 2060 CAGATGATCAAAATCAATTTTAATCAATGATTAATTAATTAATTAATTAATTAATTAATTCCT 2119  
Db 4708 TATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 4651  
Qy 2120 CTATGCAACCAAAATAGAGAAAAACAAAAATTAGAACTATCCAAAACAAAAATTAATACAT 2179  
Db 4650 ACAATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 4591  
Qy 2180 TTTTCACAAA 2189  
Db 4590 ATTACACAAA 4581

Search completed: December 11, 2005, 14:08:18  
Job time : 1805 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 10, 2005, 18:16:09 ; Search time 397 Seconds  
(without alignments)  
9886.287 Million cell updates/sec

Title: US-10-783-417-1

Perfect score: 2208

Sequence: 1 atgacatacaataacagataa.....atcatatacaaacacttga 2208

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 1303057 seqs, 888780828 residues

Total number of hits satisfying chosen parameters: 2606114

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

- Issued Patents NA:\*
- 1: /cgn2\_6/ptodata/1/ina/1 COMB.seq:\*
  - 2: /cgn2\_6/ptodata/1/ina/5 COMB.seq:\*
  - 3: /cgn2\_6/ptodata/1/ina/6A COMB.seq:\*
  - 4: /cgn2\_6/ptodata/1/ina/6B COMB.seq:\*
  - 5: /cgn2\_6/ptodata/1/ina/H COMB.seq:\*
  - 6: /cgn2\_6/ptodata/1/ina/PCTUS COMB.seq:\*
  - 7: /cgn2\_6/ptodata/1/ina/PP COMB.seq:\*
  - 8: /cgn2\_6/ptodata/1/ina/RE COMB.seq:\*
  - 9: /cgn2\_6/ptodata/1/ina/backfile1.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	340.6	15.4	3543	3	US-09-224-024-27
2	340.6	15.4	3543	6	PCT-US94-07902-27
3	188.2	8.5	2061	3	US-09-224-024-30
4	188.2	8.5	2061	6	PCT-US94-07902-30
5	98.4	4.5	3504	3	US-10-089-678-2
6	98.4	4.5	3504	3	US-10-089-678-3
7	85.8	3.9	3444	2	US-08-349-867-22
8	85.8	3.9	3444	2	US-08-349-867-25
9	85.8	3.9	3444	2	US-08-239-476-22
10	85.8	3.9	3444	2	US-08-239-476-26
11	85.8	3.9	3444	2	US-08-598-305A-22
12	85.8	3.9	3444	2	US-08-598-305A-26
13	85.8	3.9	3444	2	US-08-639-923A-22
14	85.8	3.9	3444	2	US-08-639-923A-26
15	85.8	3.9	3444	6	PCT-US95-05431-22
16	85.8	3.9	3444	6	PCT-US95-05431-26
17	85.8	3.9	3450	2	US-08-349-867-20
18	85.8	3.9	3450	2	US-08-239-476-20
19	85.8	3.9	3450	2	US-08-598-305A-20
20	85.8	3.9	3450	6	PCT-US95-05431-20
21	85.8	3.9	3450	6	PCT-US95-05431-20
22	85.8	3.9	3522	2	US-07-828-788A-9
23	85.8	3.9	3522	2	US-08-349-867-24
24	85.8	3.9	3522	2	US-08-349-867-28

25	85.8	3.9	3522	2	US-08-239-476-24	Sequence 24, Appl
26	85.8	3.9	3522	2	US-08-239-476-28	Sequence 28, Appl
27	85.8	3.9	3522	2	US-08-356-034-7	Sequence 7, Appl
28	85.8	3.9	3522	2	US-08-598-305A-24	Sequence 24, Appl
29	85.8	3.9	3522	2	US-08-598-305A-28	Sequence 28, Appl
30	85.8	3.9	3522	2	US-08-639-923A-28	Sequence 28, Appl
31	85.8	3.9	3522	2	US-08-639-923A-24	Sequence 24, Appl
32	85.8	3.9	3522	3	US-08-933-891-7	Sequence 7, Appl
33	85.8	3.9	3522	3	US-09-178-252-5	Sequence 5, Appl
34	85.8	3.9	3522	3	US-09-521-344-7	Sequence 7, Appl
35	85.8	3.9	3522	3	US-09-826-660-5	Sequence 5, Appl
36	85.8	3.9	3522	3	US-09-837-961A-7	Sequence 7, Appl
37	85.8	3.9	3522	6	PCT-US92-11337-9	Sequence 9, Appl
38	85.8	3.9	3522	6	PCT-US95-05431-24	Sequence 24, Appl
39	85.8	3.9	3522	6	PCT-US95-05431-28	Sequence 28, Appl
40	85.8	3.9	3522	9	5188960-7	Patent No. 5188960
41	85.8	3.9	4020	6	PCT-US91-02560-1	Sequence 1, Appl
42	76.4	3.5	3501	2	US-08-448-170-5	Sequence 5, Appl
43	76.4	3.5	3501	3	US-08-961-803-4	Sequence 4, Appl
44	76.4	3.5	3507	3	US-09-661-322A-21	Sequence 21, Appl
45	76.2	3.5	3504	2	US-08-291-368-3	Sequence 3, Appl

# ALIGNMENTS

RESULT 1  
US-09-224-024-27  
Sequence 27, Application US/09224024  
Patent No. 6056953  
GENERAL INFORMATION:  
APPLICANT: Leslie Hickie  
TITLE OF INVENTION: Materials and Methods for the Control of  
NUMBER OF SEQUENCES: 31  
CORRESPONDENCE ADDRESSES:  
ADDRESS: David R. Saliwanchik  
STREET: 2421 N.W. 41st Street, Suite A-1  
CITY: Gainesville  
STATE: FL  
COUNTRY: USA  
ZIP: 32606  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/224,024  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/856,226  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Saliwanchik, David R.  
REGISTRATION NUMBER: 31,794  
REFERENCE/DOCKET NUMBER: NA79  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 904-375-8100  
TELEFAX: 904-372-5800  
INFORMATION FOR SEQ ID NO: 27:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 3543 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
US-09-224-024-27  
Query Match 15.4%; Score 340.6; DB 3; Length 3543;  
Best Local Similarity 53.1%; Pred. No. 5,56-59;





Db 1263 TTTTATTAATGTCATAGCTAGATTAATTAATCTAAATGATTAATTAATTTAG 1322  
Qy 1389 TCAATTAATGATTAATTAATTAATGCTCTAACAACAACACTCAATATTCAGAGAG 1448  
Db 1323 TAAATGATTTTATTAATTAATGATTAATGATTAATGATTAATGATTAATGATTA 1382  
Qy 1449 GTCT-----TTATCTAATTAATCAAAACAACATTTTCAATTTCTAGAAAAAGA 1502  
Db 1383 ATCTGGGCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1442  
Qy 1503 CTGCAATCTAGTAAATGATTAATGATTAATGATTAATGATTAATGATTAATGAT 1562  
Db 1443 AAGAGAGATTAATGATTAATGATTAATGATTAATGATTAATGATTAATGATTA 1502  
Qy 1563 ATCCCATTTTCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1622  
Db 1503 ATC-----ATTTATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1545  
Qy 1623 TACAGGTGTA--TTAGATGACACACAGTATGTTGATTAATTAATTAATTAATTA 1680  
Db 1546 CAAGTATTAATGATTAATGATTAATGATTAATGATTAATGATTAATGATTAAT 1605  
Qy 1681 AAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1740  
Db 1606 CATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1665  
Qy 1741 ATTAAGAGCTGCTGATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1800  
Db 1666 GTTCAAGAGCTGCTGATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1719  
Qy 1801 GAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1860  
Db 1720 AAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1779  
Qy 1861 ACAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1920  
Db 1780 TCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1836  
Qy 1921 ATACCACTCAAGATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1980  
Db 1837 GAATGAGGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1896  
Qy 1981 GGAATTTTGGGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2040  
Db 1897 AAAGATTTTGGGATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1956  
Qy 2041 CCATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2097  
Db 1957 TCTCTGTTGTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2016  
Qy 2098 GAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2157  
Db 2017 GAATTTTGGGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2076  
Qy 2158 ATCCAAACAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2200  
Db 2077 GTACAAACAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2119

## RESULT 3

US-09-224-024-30  
Sequence 30, Application US/09224024  
Patent No. 6056953

## GENERAL INFORMATION:

APPLICANT: Leslie Hickie  
APPLICANT: Jewel Payne  
TITLE OF INVENTION: Materials and Methods for the Control of  
NUMBER OF SEQUENCES: 31  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: David R. Saliwanchik  
STREET: 2421 N.W. 41st Street, Suite A-1  
CITY: Gainesville

STATE: FL  
COUNTRY: USA  
ZIP: 32606  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/224,024  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/856,226  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Saliwanchik, David R.  
REGISTRATION NUMBER: 31,794  
REFERENCE/DOCKET NUMBER: MA79  
TELEPHONE: 904-375-8100  
TELEFAX: 904-372-5800  
INFORMATION FOR SEQ. ID NO.: 30:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2061 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
US-09-224-024-30

Query Match 8.5%; Score 188.2; DB 3; Length 2061;  
Best Local Similarity 50.0%; Pred. No. 1.6e-28;  
Matches 1107; Conservative 0; Mismatches 923; Indels 185; Gaps 18;

Qy 1 ATGAATCAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 60  
Db 1 ATGAATCAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 54  
Qy 61 TTTCGGAACGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 120  
Db 55 GGTTTAGCAAGCTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 114  
Qy 121 TTACAAAACGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 180  
Db 115 CTGAAAAACGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 174  
Qy 181 GATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 240  
Db 175 AATTAATGCGGGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 234  
Qy 241 GTATCCGTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 300  
Db 235 GTATGAGGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 270  
Qy 301 ATTAATGCTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 360  
Db 271 GTCTAGCTGAGTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 330  
Qy 361 GAACAAACGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 420  
Db 331 GACCTG---CAATGTTTGGAGGATTAATTAATTAATTAATTAATTAATTAATTA 387  
Qy 421 CCGTTAACAAGGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 480  
Db 388 ATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 441  
Qy 481 TTCAAGGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 540  
Db 442 CAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 501  
Qy 541 GAATTAACCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 600



Db 502 GCTAAGACAGTA - CAGTATCTCTTACTACTAGAACCTATATATAGATTAAGATTAGA 560  
 Qy 601 AATGTTCAATGATTTTATGAGAAATACCGTTCCACTGAACTTATATAAGC 660  
 Db 561 TATGTTAAAA-----ATATGCTAGCTATGCAATA 591  
 Qy 661 CTATTAATACCTATTTATGCGAAGCTGCTAAATTTTCAATTAATTTATTAACAAGCT 720  
 Db 592 CCAACACTCCCTGCAATGACAAATAGCTACTGGCACTTAATTTATTAACAAGCT 651  
 Qy 721 GCTGAATTTGCTGATGATGCAATGCAATATACCTCTTCAAAATGCACTTAATGCT 780  
 Db 652 GCTACCTATTAACAAATATAGCTGCTGCAAAATCAAGGAT-----AATTCAGATCTTTC 705  
 Qy 781 GGAACATGAGATGATATATTAATTTAAAGAAATATACCTTAATATATGATATAT 840  
 Db 706 AATTAATCTATTAATCTATGAGGCTATTTAAACGTAATATCAAGAAATATATCTGATAT 765  
 Qy 841 TGTGCAATACCTATATAGAACAGACTAAAAATCTTAGACGAACCAATATGAAATG 900  
 Db 766 TGTATACAAAGCTAGACAGTACGACTAATATGATTAAGAACTAATATACCAAGCAATG 825  
 Qy 901 AGTATATTTATGATCTATGCAATATATGACCAATATCTGATATGATACATCTCTCA 960  
 Db 826 AATATGATATATATCTTACCGTTAGAAATGACTTAACTGTGTGATCTTATTCATAT 885  
 Qy 961 TTTTCTTATATGATATTAATAAGATATATGAGATTCATAGAGAAATAGAAATTAAGC 1020  
 Db 886 TTTCAATTTATGACCCGAAAAATATCC-----AATAGCA 921  
 Qy 1021 AATTAAGATGATGCAAGAAATTTTATACACTGA -AATTAATTTGATGCTCTCC 1079  
 Db 922 GTTAATCTGAACTTACAGAGAAATTTATGAAATTTTATTCAGATATCTTGAAC 981  
 Qy 1080 TCACTTATGATGCAACCAATCTATGATGATGCAATATTAATTAACGCTGCAATTT 1139  
 Db 982 ATTAACAGACTGAAAAATGATATTAATGAAATCTCAATTTTATCTTGATTAACCA 1041  
 Qy 1140 TAAATTTATGATTTTATGAAACATTTATTTTATACAGAAATTAACAAATTTGGGAA 1199  
 Db 1042 GGGCGTTTATACAGAAATTTCTGAGCAATCTTGATCTTATGATATTTTCTTTT 1101  
 Qy 1200 TCGTTTATGTTTATTTTCTAATCTGATGACACTTATTAACATATATATGAAAC 1259  
 Db 1102 ACAGGTAACAG -ATGCGCTTTACACATATCTAATATATGCAATATCTGGGAGC 1160  
 Qy 1260 TTTATATGAGAAAGAACAGTTTCAACCAACAAACAAATTAAGACATTTGAAATCTTA 1319  
 Db 1161 GGTTCATGGAATATTTATTTCTCAAGACATCCAAA-----GTAATTTCC 1205  
 Qy 1320 TAAATTTCAATTTGATGATGACATCACTCTCTGTTTCCCTATTTCAACCACTT 1379  
 Db 1206 TTTTATATGAAACAAACCTATGATTAAGTTCGAAATTTGACAGACATGAGACTACTCAGA 1265  
 Qy 1380 TATAATTAATCAATTTGAACTTTATTAATGCTCATCTAACAACACTCAATATTTT 1439  
 Db 1266 TATAATTAATGAAATGATTTTTCGAATAGC-----AGTGAATTTTCGATATTT 1319  
 Qy 1440 AGCAGAGAGGCTTTATCTAATTAATCAAAACAACTTTTTCATTTCTTGAAGAAAA 1499  
 Db 1320 ATCCAAATTCACAAATAGAAATTAATTAAGAACTGATCTTATATGATTTCAAAACA 1379  
 Qy 1500 AGACTGCAATCTAGTTATGATCCAGTTTTCACCAAACTTAATTAATGATGATAT 1559  
 Db 1380 AACATG-----GAAAAATTAAGAAATATGCTATAC 1409  
 Qy 1560 TTTATCCATTTTCTATTTACTTATTTCTATGATGATTTGATTAACAGCTCAATAT 1619  
 Db 1410 TCTATCTGATATTAATACTGATATAT-----ATATTTCTAGTTATGAGAAAG 1460  
 Qy 1620 AGATACAGGTGATTAAGATGACACACAGTATGTTGATGATATATATGCAATATCAGA 1679  
 Db 1461 AAGAAGAGTTGATTTAGTTGACACATCTAGTGTGATTTCCAAATTAACATATGATTT 1520

Qy 1680 TAAATTAATTAATGATATCCAGCAATCAAGATTAACATCTTGATACAACTTAAGT 1739  
 Db 1521 AGATTAATCAATCCCAATTCACAGCTCTTAAGAACTTTGAAGTATCTGATTCGAAAT 1580  
 Qy 1740 AATTAAGAGCTCTGCTATACAGAGAACTTGTATTAATTAACAAGTCAAGGCTTT 1799  
 Db 1581 TGTGAAGGTCTGTGCTACACAGGTGAGACTTGTATTTCTTAAGATATGATGATTT 1640  
 Qy 1800 AGAAATTAATGTAAGTATCTTAATTTCTTACCAATCTTATTTCAATTAAGCTGATATGC 1859  
 Db 1641 TAGAGTTGATTTTAAAAAATGTTTCTGACAAATATCAAGTACGTA---TTGTTATGC 1697  
 Qy 1860 TACAATGCTCTGGAATATCTCTTAATATATCTCTTAACATCAAGGATATAG 1919  
 Db 1698 TACTATCTCTCAAAAGACAGAGT--ATCTTAACCGGAATATGATATCTATAGTGTGAG 1755  
 Qy 1920 AATACCACTCAAGCTCAACACACTTTTCTGTGTAACAAATTAATTAATTAACAATA 1979  
 Db 1756 CTCCCTAGTACCACTTCCGCAAAAC-----CAATGCTACAGATTTTAACTA 1805  
 Qy 1980 CGGATTTTGGGATTTTCCAAATTTCCAGTACGT-----ACATTAACCTTTAA 2030  
 Db 1806 TCGAGATTTGATATGATATGATATGATTTCCAGAACAGTTCCAAATTAACATTTGAAGAGA 1865  
 Qy 2031 TCGAAACATACATTTATTAATTTAATCGGCAAGATGATCAATTTTATATCATTTGA 2090  
 Db 1866 AGACTTTATTAATGACTTATATGATATGATACCAAAATATTCATATATATATATATGA 1925  
 Qy 2091 TAAATTAATTTATACCAATTTACTCTCTATGACCAAAATGAGAAAAAATAATTT 2150  
 Db 1926 CAAATGCAATTTATTCAAATCACTCAATCTGATATGATATATACAGAGAAACAAATAT 1985  
 Qy 2151 AGAACTATCCAAACAAATTAATTAATTTTTCACAAATCATACAAACACTT 2205  
 Db 1986 AGAAAAACACAGAAATGAGATTTATTTGTTAATTAACAAAGTTCTT 2040

RESULT 4  
 PCT-US94-07902-30  
 ; Sequence 30, Application PC/TUS9407902  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Street address: 4980 Carroll Canyon Road  
 ; APPLICANT: City: San Diego  
 ; APPLICANT: State/Province: California  
 ; APPLICANT: Country: US  
 ; APPLICANT: Postal code/Zip: 92121  
 ; APPLICANT: Phone number: (619) 453-8030 Fax number: (619) 453-6991  
 ; APPLICANT: Telex number:  
 ; TITLE OF INVENTION: Materials and Methods for the Control of  
 ; NUMBER OF SEQUENCES: 31  
 ; CORRESPONDENCE ADDRESSES:  
 ; ADDRESSEE: David R. Saliwanchik  
 ; STREET: 2421 N.W. 41st Street, Suite A-1  
 ; CITY: Gainesville  
 ; STATE: FL  
 ; COUNTRY: USA  
 ; ZIP: 32606  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: Patent Release #1.0, Version #1.25  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: PCT/US94/07902  
 ; FILING DATE:  
 ; CLASSIFICATION:  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Saliwanchik, David R.  
 ; REGISTRATION NUMBER: 31,794  
 ; REFERENCE/DOCKET NUMBER: MA79

TELECOMMUNICATION INFORMATION:  
TELEPHONE: 904-375-8100  
TELEFAX: 904-372-5800  
INFORMATION FOR SEQ ID NO: 30:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2061 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
PCT-US94-07902-30

Query Match 8.5%; Score 188.2; DB 6; Length 2061;

Best Local Similarity 50.0%; Pred. No. 1.6e-28; Matches 1107; Conservative 0; Mismatches 923; Indels 185; Gaps 18;

1 ATGATATCAAAATTAACGATATATACGAAATATGAAATTTATGATTCGATACCTCACTTAT 60  
1 ATGATATCAATTCAAAATTAAGAAATGAAATATGAAATATTCATGCTC-----CATCCAT 54  
61 TTTCCGAAACAGAAACAGTAATGATTCTAGATACCTTACACAAATATCCAAATCAACCA 120  
55 GATTTTACCAAGTCTAATATACATCTATCTAGATATCCATTAGCAAAATPAGCCAAATCAACCA 114  
121 TTACAAAACAAAATTAAGAAAGTGGCTCAATATGTCTCAAGGAAATACACATATGT 180  
115 CTGAAAAACAGAAATTAACAAAGATTGGCTCAATGTGTCAAGATATATCAACAAATATGGC 174  
181 GATAATTTGAGAGACATTTGCTAGTCTGATACAAATTTGCTGACAGTTAGTGCAGTACTAT 240  
175 AATATATGGGGGAATTTTGTATAGTTCTGAAACATATGTTGAGATTTAGTGCAGTATAT 234  
241 GTATCCGTAATCTGTATGACCGGATATAGTGGGCTCACTTCAATATCCGACCGATAGGA 300  
235 GTAGTAGACACTATGTTAGAGCTTTGCTGCTCT----- 270  
301 AATAATAGGTCATATATATATCTTTTGGTACCCCAATCACGTCTTTTGGCCGCGGA 360  
271 GCTTAGCTGACAGTATATATATCTTTTGGACTTTGTGCGCATCTTTTGGCAAGATCT 330  
361 GAACAAAGCAAAACAGTATGACACAAATTTATTAATGAGAAATTTTGTGATACA 420  
331 GACCTG---CAATGTTTGGCAGAGATTGTTAAACATCGAGAGAGGCCCTATACAGAA 387  
421 CCGTTACAGAAAGCATATAAACAGTAAAGTTACAACTTTAGAGAGATTAGCAAAATA 480  
388 ATAGATATAAAACATATATATATGTACTACTT-----CTATCGTACACCTATATAAAAT 441  
481 TTACAAAGCTATATATACGACATATAGATGAGAGAAATTTAAAGCTACAGCTCCT 540  
442 CAACCTGATTAATATCAAGAAATTTTGGATTAATGAGGCGCAGCAGTACACAGCTAT 501  
541 GAATTAACACATCATCATGACATTTACACAGAGCTGCTTACCTTTAAATATGCAATTTAG 600  
502 GCTAAGAGAGTA-CATGATCTCTTACTACCTTAGAACCTATATATGATTAAGATTTTGA 560  
601 AATGTTCACAATGATTTTATTCGAGAAATACCTGTTTCAACTTGAACCTTATATAACG 660  
561 TATGTTAAAA-----ATAATGCTAGCTATGAAATA 591  
661 CTATTTACTACCTATTTATGCGCAGAGCTCTAATTTTCAATTTTAAATTTATACACAGGT 720  
592 CCAACACCTCCCTGATATGACAAATATGCTACTTGGCACTTGAATTTATTAACATGCT 651  
721 GCTGATTTGGCTGATGATGAAATGACATATACATCTTCAACAAATGGAAGCTAATGCT 780  
652 GCTACCTATTAATATATATGCTGCAAAATCAAGTAT-----AAATCCAGATCTTTC 705  
781 GGAACATGAGTACTATTTATTAACCTTTTAAAGAAATATATACCTTAATATAGTAACTAT 840  
706 AATTCATCTAATTAATCTATCAGGGCTAATTTAAACGTAATAATACAAAGATATAGCATAT 765  
841 TGTGAAATACCTATATGAACAGACTAAAAATCTTAGAGAGAAACCAATATGAAATGG 900

766 TGTATACAAACGTACATGACAGGCTACTATGATTTGAATCTATATCTAATACCAACATGG 825  
901 AGTATATTTAATGACATATCAAGATATATAGCAATTAATCTGATATTAATACATCTCTCAA 960  
826 AATATGTAATATCTTACCGTTAGAAATGACTCTAATCTGTGATTAATCTTATTTGCTATT 885  
961 TTTTCTTATATGATATATAAAGATATAGAGATTCAATAGAGGAATAGAAATGAAAGGC 1020  
886 TTTCCAAATTAATGACCCGAAATAATATCC-----AATGGA 921  
1021 ATTAGAATGAACTCAACAGAGAAATTTATACACGTA-AATAATTTGATGCTCTCC 1079  
922 GTTAAATCTGAATCTACAGAGAAATTTATACGAATTTATATCAATATCTGATATAGAAC 981  
1080 TCAATTAAGATTCACACCAATCTAGCTAGAGATATATATTTTAAACAGTGCAGATT 1139  
982 ATTAACAACTAGAAATATGATTTAATCTAGAAATCTGATATTTTACTGTGATTAACCA 1041  
1140 TAAATTAATTTCAATTTTAAACAATTTATTTTATACAGAAATATCAAAATTTGGGAA 1199  
1042 GGGCGTTTATACAGAAATTTCTGAGACATTTGATCTTATGATATATTTTCTTTT 1101  
1200 TCGTTTATGTTGATTTCTAATGCTGATGCACTTATAGCAATCTAATCTGAAC 1259  
1102 ACAGTAAACAG-ATGGCTTTTACACATCTAATATGATGACAACTAATCTGGGAGC 1160  
1260 TTTATATGAGAAAGAACAGGTTTCAACCCACAAACAAATTAAGACATTTGATCTTA 1319  
1161 GGTTCATGACATATATATTTCTCAAGACATCCAA-----GATTTTC 1205  
1320 TAAAGTTCAATTTGATATGATATGACATACCTCTGTTTCCCTTATTAACCAACACT 1379  
1206 TTTTATAGAAACAAACCTATGATAGTGAATTTGACATGAGAGATGATCTGAA 1265  
1380 TATATATATCAATGAATGAATTTTAAATGCTACTTAACCAACACTCTAAATATTC 1439  
1266 TATATATATGAATATGATTTTTCGAAATAGC-----AGTGAATATTTGATATTC 1319  
1440 AGCAGAGGCTCTTATCTAATATACAAACACAACTTTTTCATTTCTAGAAAAA 1499  
1320 ATCCAAATTAACATATAAATATATTAAGAACTGATCTTATATGATCCAAATCA 1379  
1500 AGACTGCAATCTATGATATGATCCAGGTGTTTCAACCACTTATATATCTATATCTAT 1559  
1380 AACATG-----GAAAAATTAAGAAATATGCTCATAC 1409  
1560 TTTATCCATTTTCAATTTTACTATTTCTATGTTGATGATTAAGCTACCAATATAT 1619  
1410 TCTATGCTATATATAAATCTGATATAT-----ATATTTCACTAGTTAGAGAAAG 1460  
1620 AGATACAGTGTATTAAGATGACACACAGTATGTTGATATATATATGCAATATCAGA 1679  
1461 AAGAAAGTGTGCAATTTAGTTGACACATACATAGTGTGATTTCCAAATATCAATATGAT 1520  
1680 TAAATATATTAATCAATGATCCAGCAATCAAGGTATCAATCTTGATACAACTTAAGT 1739  
1521 AGATTAACATACCCAAATCCACGCTTAAAGGCTTGAAGGTATGATCTGATTCGAAAT 1580  
1740 AATTAAGAGCTGCTGATACAGAGAGAACTTGTTATTTTCAAGTCAAGGCGCTT 1799  
1581 TGTGAAAGGCTGCTGATACAGAGTGAAGCTTGATTTCTTAAGATATGATGATTT 1640  
1800 AGAATATCATGTAATCTCTAATTTTACACATCTTATTTCAATTTAGACTGCTATGCT 1859  
1641 TAGAGTATGATTTTAAATAATGTTCTGCAATATCAAGTACGTA---TTGTTATGCT 1697  
1860 TACAAATGCTGCGAAATATCTCTCTAATATATCTCTTAACATATACAGAGATATAGG 1919  
1698 TACTAATGCTCCAAAGCAACAGT--ATTCTTAACCGGAATAGATATCTAATAGGTGAG 1755  
1920 AATTCACCTCAAGCACTCAACCAACTTTTCTGATGAGAAATTAATATTAATTTACATA 1979

Db 1756 CTCCTAGTACCACTCCGCCAAAAAC-----CAAACTGCTACAGATTTAACAATA 1805  
Qy 1980 CGAGATTTTGGGTATTTCCAAATTTCCAGTACAGT-----ACATTCCTTTAAA 2030  
Db 1806 TGCAATTTTGGATTTGTAACATTTCCAGAAACAGTCCAAATTAACATTTGAAGAGA 1865  
Qy 2031 TCGAAACATACCAATTTATTTATTTGTCGAGATGATCAATTCATTTTAATCAATGA 2090  
Db 1866 AGACATTTATTAATGACCTTATATGATGACCAAAATCATTCATATATATATATTGA 1925  
Qy 2091 TAAATTTGAATTTATACCAATTTCTCTATGACCAAAATAGAGAAAAAATAATT 2150  
Db 1926 CAAAATCGAATTTATTCACATCTCATCTGATTTAGATTTTACAGAGAACAAATAT 1985  
Qy 2151 AGAACTATCCAAACAAATAATATACATTTTCCAAATCAGTACAAAAACACTT 2205  
Db 1986 AGAAAAACACAGAAATATGTAATGATTTATTTGTTAATTAACAAAGTTCTT 2040

RESULT 5  
US-10-089-678-2  
; Sequence 2, Application US/10089678  
; Patent No. 6962977  
; GENERAL INFORMATION:  
; APPLICANT: ASANO, Shinichiro  
; TITLE OF INVENTION: PROTEIN INVOLVED IN INSECTICIDAL ACTIVITY, DNA ENCODING THE PROTEIN, A  
; FILE REFERENCE: 068821  
; CURRENT APPLICATION NUMBER: US/10/089,678  
; PRIOR FILING DATE: 2002-04-03  
; PRIOR APPLICATION NUMBER: JP 2000-236140  
; PRIOR FILING DATE: 2000-08-03  
; PRIOR APPLICATION NUMBER: PCT/JP01/06660  
; NUMBER OF SEQ ID NOS: 3  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 2  
; LENGTH: 3504  
; TYPE: DNA  
; ORGANISM: Bacillus thuringiensis  
; FEATURE:  
; NAME/KEY: exon  
; LOCATION: (1)..(3501)  
; OTHER INFORMATION:  
US-10-089-678-2

Query Match 4.5%; Score 98.4; DB 3; Length 3504;  
Best Local Similarity 43.4%; Pred. No. 1.6e-10;  
Matches 602; Conservative 0; Mismatches 751; Indels 33; Gaps 2;

Qy 612 TGATTTTATTCAGAAATACCTGTTTCCAACTTGAACTTATTAACGCTATTACTAC 671  
Db 531 TGATTTTGTGCTAAATCCATCTTTGCAATATCTGACAGAAATACATTTATATC 590  
Qy 672 TATTTATGCGAAGCTGCTAATTTTCAATTAATTATTAACAAGGCTGAAATGGC 731  
Db 591 AGTATGACAGCAAGCAATTTTCAATTTGCTATTTATACAGATGCTTCCATTTTGG 650  
Qy 732 TGATGAATGGAATGAGATATATCATCTTACAAATGAACTTAATGCTGAACATCAGA 791  
Db 651 AGCAGAGTGG-----GGATTCAACACCGAGAGAAATTTTC 683  
Qy 792 TGACATTTATTAACCTTTTAAAGAAATATACCTAATATTAATTAATGTCGAAATTC 851  
Db 684 CACATTTTATGATCGTCAGGTGACAGTACCGCCCAATATCTCGATTTATGTAAGTG 743  
Qy 852 CTATGAGACAGACTPAAAAATCTTAGAGACGACCAATATGAATGAGATATTTTAA 911  
Db 744 GATTAACACTGCTTATGATTAATTAAGTACGATGTCGAAATGCTGTAAGATCA 803  
Qy 912 TGACATGAGAGATATATGACATTTACTGATTTAGATTCATCTCTCAATTTTCTTTATA 971  
Db 804 CCAATTTCCGAAAGAAATGACATTTACTGATTTAGATTTAGAGGTTATTTCCAAACTA 863

Qy 972 TGATATTAAGATATAGATTCATATAGAGAAATAGAAATTAAGGCAATTAAGATGA 1031  
Db 864 TGACACAGTATCGATCCAAATCGAAACAGGCCCACTTACAGGGAAGTATATACGA 923  
Qy 1032 ACTCAGAGAGAAATTTATACACTGAAATTAATTTTGAATGCTCTTCTCAACTTAGAGT 1091  
Db 924 TCCATATGATATTTTAAACAGAGAAACAGTGTGATTTTGTGGCGTGTGATCACTTAACG 983  
Qy 1092 TCAACCAATCTAGCTAGATGGAATATATTAATTAACAGTGCAGTTTAAATATTTTC 1151  
Db 984 TGATATTTCTTTTTCAGAGTGAAGCGCTGTAATTTGTCACCACTTATTTGATAT 1043  
Qy 1152 ATTTTGAACAATTTATTTTATTTATACAGAAATTAACAATTTGCGAATCTTTAGTTGG 1211  
Db 1044 ACTCAGTAATTAATGATTTTATACAGAGAGCGGGCTTCCCTGATTAATAGGAATA 1103  
Qy 1212 TATTTCTAATGCTGATGACCTACT-----TATGCAATATCTTATGTAACCTTATA 1265  
Db 1104 CTTGAATATTTGGGATGAGACATTCATTAATTAATTAATTAATGCAATGCTCATCAGCAT 1163  
Qy 1266 TGGAGAAAGAACAGTTCCACCCACAAACAAATTAAGACATTTGATCTTATTAAGT 1325  
Db 1164 AGAAGTATTAACGATACATTTACTTTCACAAATTCAGATTTATGACAAATPA 1223  
Qy 1326 TTCAATTTGTAATGATAGACATCACTCTGTTTCCCTTATTCACACCACTTTATAT 1385  
Db 1224 GATATATCTTTGAGGTGATGATTAAGGGCGGATTTAGCTATTAATACGACAGATATA 1283  
Qy 1386 TAATCAATTAATGACTTTTAAATGCTCATTAACAACACTCAAAATATTCGACAG 1445  
Db 1284 TGAAGTCCGATACGCTATTTTAACTGCTGACAGAAATCAGAGATCAGATTCGTTGG 1343  
Qy 1446 AGGCTTATATCTAATTTATCAAAACAACTTTTTCATTTCTAGAAAAAAGACGTG 1505  
Db 1344 AGGTTTACGTACTCAAAACCAATACATCATGCAATGATGTAACAAATTAACAATTC 1403  
Qy 1506 CAATCTAGTATTTGATCCAGGTTGTTCAACAACTTTATTAATCTATGCTAATTTATC 1565  
Db 1404 GATGATGAATATCCCTCCAGAGAAATGAGCCACTTAGTAGAGGATTTACCATAGATATTC 1463  
Qy 1566 CCAATTTTCAATTAATTTACTTATTCCTATGATGATGATTAACGCTTACAAATATAGATTC 1625  
Db 1464 TCATATACCTCTTATTTCTTTTCTAAGAAATGATAGTGTCTGCTGATATGCAATCT 1523  
Qy 1626 AGGTATATTAAGATGACACAGTATGTTGATGATTAATATGCAATATGCAATTAAT 1685  
Db 1524 CCTGTATTTGCTTGACACATCGAGTGGAGATGTTACAAATACAGTTTATTCAGATPA 1583  
Qy 1686 AATTAACATGATCCAGCAATCAAGGTAACATCTTGATCAAACTCTAAGGTAAATGA 1745  
Db 1584 AATTAACATGATCCAGGTTGTAAGGCAATCTTTAGTTTCAAGTATCTACGTTATTTA 1643  
Qy 1746 AGACCTGTCATACAGAGGAAATCTGGTTATTTTCAAAAGTCAAGGCGTTAGAAAT 1805  
Db 1644 AGGCTTGTGATTTACAGAGCAATATCTTTTAAAGAACAGTAGTGTCCGTTAGCTTA 1703  
Qy 1806 TACATGTAATCTCTAATTTTACCAATCTTATTTTCAATGATGCTTGCATATGCTACAA 1865  
Db 1704 TACTAGTGTCTGTAAATACCAATTAATCAAAAGATATGTCGMAAATATACGTTATGC 1763  
Qy 1866 TGGTGTGGAATATCTTTCCAAATATATCTTTTACAAATACAGAGATATAGAAATCC 1925  
Db 1764 TTCTACTACTAATCTTACGATTTTGTATACAAATTTTGGAACTGCAATTTACTCTATAAA 1823  
Qy 1926 ACTCAACAGACTCAACAACTTTTCTGTACAAATTAATTAATTTTACAAATACGAGA 1985  
Db 1824 TGTATATTAACATGATTAAGGAGATGATTTAACTTAATATGACTTAGCAAC 1883  
Qy 1986 TTTTGG 1991  
Db 1884 TATTTGG 1889

RESULT 6  
US-10-089-678-3  
; Sequence 3, Application US/10089678  
; Patent No. 6962977  
; GENERAL INFORMATION:  
; APPLICANT: ASANO, Shintaro  
; TITLE OF INVENTION: PROTEIN HAVING INSECTICIDAL ACTIVITY, DNA ENCODING THE PROTEIN,  
; FILE OF INVENTION: NOXIOUS ORGANISM-CONTROLLING AGENT AND METHOD  
; CURRENT APPLICATION NUMBER: US/10/089,678  
; CURRENT FILING DATE: 2002-04-03  
; PRIOR APPLICATION NUMBER: JP 2000-236140  
; PRIOR FILING DATE: 2000-08-03  
; PRIOR APPLICATION NUMBER: PCT/JP01/06660  
; NUMBER OF SEQ ID NOS: 3  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 3  
; LENGTH: 3690  
; TYPE: DNA  
; ORGANISM: Bacillus thuringiensis  
US-10-089-678-3

Query Match 4.5%; Score 98.4; DB 3; Length 3690;  
Best Local Similarity 43.4%; Pred. No. 1.6e-10;

Matches 602; Conservative 0; Mismatches 751; Indels 33; Gaps 2;

QY 612 TGATTTATTCGAGAAATACCTGTTCCAACTTGAACTTAATAAGCTATTACTAC 671  
DB 717 TGATTTGTGTGTAATAATCCATCTTTGCAATATCTGACAGAGAACATTTATTC 776  
QY 672 TATTTATGCGCAAGCTGCTAATTTTCATTTAAATTTATTAACAAGGTGCTGAATGGC 731  
DB 777 AGTGATGACACAGCAGCAATTTACATTTGCTATTATTACAGAGTCTTCATTTTGG 836  
QY 732 TGATGAATGGAATGAGATATACATCCCTTCAAAATTTGAACCTTAATGCTGAACATCGA 791  
DB 837 AGCAGAGTGG-----GGAATTCACACAGAGAAATTTTC 869  
QY 792 TGACTATTATAAATCTTTTAAAGAAATATATACCTTAATATAGTATCTATTGCAAAATAC 851  
DB 870 CACATTTATGATGCTGAGTGACGTCACGCCCAATATCTCGATTTATTTGTTAAAGTG 929  
QY 852 CTATGAAACAGACTTAAATAATCTTGAAGACGAAACCAATATGAATGAGATATATTTAA 911  
DB 930 GTATTAACACTGCTTAGATTAATTAAGTACGAATGCTCAAGTGGCTGAAGTATCA 989  
QY 912 TGACATGAGAGATATATGACCATCTATGATATACATCTCTCAATTTTCTTATA 971  
DB 990 CCAATTCGAAAGAAATGACATTAAGTATTAATTTAGATTTAGATTTTCCAAACTA 1049  
QY 972 TGATATAAAGATATAGATATTCATAGAGAGATAGAAATTAAGGATTAAGAAATGA 1031  
DB 1050 TGACACAGATGATGATTCATGAAACAAACGCCCACTTACAGCGGAGATGATACGA 1109  
QY 1032 ACTCAAGAGAAATTTATCAACTGAATTAATTTGATGCTCTCTCACTTAAGAT 1091  
DB 1110 TCCATATGATTTTAAACAGAAACAAAGTGTGATTTTGTAGGCGTGTGATCACTTAACAG 1169  
QY 1092 TCAACCCCAATGAGTACAGTAAATATATTAACAGTCAAGTAAATTTTCTTATTTTC 1151  
DB 1170 TGAATATTTCTTTTGAAGTGAAGGCGGTGAATTTGTTTCCACACACCTTAATTTGATAT 1229  
QY 1152 ATTTTGAACAATTTATTTTATACAGAAATCAATTTTGGGAATCGTTAGTTGG 1211  
DB 1230 ACTGATGAAATGAAATTTTATACAAACAGCGGCGTTCCTTGAATATATAGGAATA 1289  
QY 1212 TATTTCTAATGCTGATGACCTACT-----TATAGCAATATCTAATCTGAACCTTTATA 1265  
DB 1290 CTTGAATATTTGGTGAAGACATCTTAATAATATTAATAATGCAATGCTCATGAGATT 1349  
QY 1266 TGGAGAAAGAACAGTTCAACCCACAAACAAATATAGACATTTGATCTTATTAAGT 1325

DB 1350 AGAAGTAAATTAACGATGATTAATCTTTAAACAAATCAAGATATTAATTTAGCAATTA 1409  
QY 1326 TTCAATTTGTAATGATGACAAATCACTCTGTTTCCCTATTTCACACACTTTATAT 1385  
DB 1410 GGAATCTTTCAGGTTGATGATCAATTAAGGCGGATTTAGCTTAATTAACGACAGATTA 1469  
QY 1386 TAATCAAAATTAATTAATTTAAATGAGCTATCAACACACTCAATTAATTTGAGAG 1445  
DB 1470 TGAAGTCCGATGATGATTTTAACTGTTGACAGAAATACAGATCAGGATCAGTTGG 1529  
QY 1446 AGGCTCTTATCTAATTAATCAAAACACACTTTTTCATTTCTTGAATTAAGAAAGCTG 1505  
DB 1530 AGGTTTACGATCAAAACCAATACATGCAAGTATGACACAAATTAATCAATAC 1589  
QY 1506 CAATCTAATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1565  
DB 1590 GATGATGAAATCCCTCAAGAGATGAGCCATTAAGAGGATGATGATGATGATGATGATGAT 1649  
QY 1566 CCAATTTTCAATTAATTAATTTCTAATGATGATGATGATGATGATGATGATGATGATGAT 1625  
DB 1650 TCAATATCACTCTAATTTCTTTTCTAAGAAATGATGATGATGATGATGATGATGATGATGAT 1709  
QY 1626 AGGATATTAAGATGACACAGATGATGATGATGATGATGATGATGATGATGATGATGAT 1685  
DB 1710 CCGTATTTGCTTGAACACATGAGATGAGATGATGATGATGATGATGATGATGATGATGAT 1769  
QY 1686 AATTAATATATCCCAATCAAAAGGTAACATCTTGAATACAACTTAAGTAAATTTGA 1745  
DB 1770 AATTAATCAATCAATCAATTTTAAAGGACATCTTATGTTTCAAGTATCTATTA 1829  
QY 1746 AGAAGCTGCTATACAGAGGAAATCTGTTTATTAACAAAGTCAAGGCGTTTGAAT 1805  
DB 1830 AGGCTGCTGATTAACAGAGGCAATATCTTAAAGAAACAAAGTATGATGATGATGATGAT 1889  
QY 1806 TACATGTAATCTCTAATTTCTAACAATCTTAATTTTCAATTTGATGATGATGATGATGAT 1865  
DB 1890 TACTAGTCTCTGTAATAATCAATTAATCAAAAGATGATGATGATGATGATGATGATGAT 1949  
QY 1866 TGGTGTGAAATATCTTCTTAATATATCTTTTCAATATCCAGAGTAAATAGGAATACC 1925  
DB 1950 TTCTACTACTAATCTTAAGATCTTTTGTAAATTTTGTGAATCTGATCTTATTA 2009  
QY 1926 ACTCAAGACATCAACAACTTTTCTGATCAAAATTAATTAATTTTAAATTAACGAGA 1985  
DB 2010 TGTAATTAACCATGATTAAGGAGATGATTAACATTTAATCAATTTGATGATGATGATGAT 2069  
QY 1986 TTTTGG 1991  
DB 2070 TATTTGG 2075

RESULT 7  
US-08-349-867-22  
; Sequence 22, Application US/08349867  
; Patent No. 5508264  
; GENERAL INFORMATION:  
; APPLICANT: Bradfisch, Gregory A.  
; APPLICANT: Thompson, Mark  
; APPLICANT: Schwab, George R.  
; TITLE OF INVENTION: No. 5508264el Pesticidal Compositions  
; NUMBER OF SEQUENCES: 34  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: David R. Saliwanchik  
; STREET: 2421 N.W. 41st Street, Suite A-1  
; CITY: Gainesville  
; STATE: FL  
; COUNTRY: USA  
; ZIP: 32606  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/349,867  
FILING DATE:  
CLASSIFICATION: 514  
ATTORNEY/AGENT INFORMATION:  
NAME: Saliwanchik, David R.  
REGISTRATION NUMBER: 31,794  
REFERENCE/DOCKET NUMBER: MA86  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 904-375-8100  
TELEFAX: 904-372-5800  
INFORMATION FOR SEQ ID NO: 22:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 3444 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
US-08-349-867-22

Query Match 3.9%; Score 85.8; DB 2; Length 3444;

Best Local Similarity 53.8%; Pred. No. 5.1e-08;  
Matches 220; Conservative 0; Mismatches 162; Indels 27; Gaps 1;

QY 589 ATACGATTGGAGAAATGTTACAAATGATTTATTCGAGAAATACCTGGTTCCAACTTGA 648  
DB 388 ATTCGATTGGCTAATACAGACAGCGCTTAATACGCAATTAATTTACACTTCA 447  
QY 649 ACTTATAAACGCTATTACTTACTTATTAATGCGCAAGCTGCTAATTTCAATTAATTA 708  
DB 448 AGTTTGAATCCCTCTTTATCGGCTATGTTCAAGCGCGCAATTTACATTTACACTA 507  
QY 709 TTACAAACAAGTGCTGAATTTGCTGATGAAATGCAATATACATCTTCACTCAAA 768  
DB 508 TTAAGACGCGTATGCTGTTGGGAGGTTGGGAGCTGATATGCTATGTAATAT 567  
QY 769 GAACCTAATGCTGAACATCAGATGACTTATTAATCTTTAAAGAAATATACCTAA 828  
DB 568 C-----ATTATATAGATTATTAATCTTTATTTATGA 600  
QY 829 TATAGTAATCTATTTGCAAAATACCTATAGAACAGACTTAAATCTTAGAGACGAACA 888  
DB 601 TATAGCAAACTTTGTTGGACATACATCAAGATTAGAAACTTAGAAGCTACTAT 660  
QY 889 AATATGAATGAGATATTTATGACTATGCAAGATATATGACCACTTACTATTAAT 948  
DB 661 ACTGACAAATGGGCAAGATTCAATCAGTTAGAGAGATTAACTTACTGTTATGAT 720  
QY 949 ACCATCTCTCAATTTCTTTATATGATATTAATAAGATATAGATTCA 997  
DB 721 ATCGTGTCTTTTCCGAACTACATGTTAGAACTATCCAAATTCAAA 769

RESULT 8  
US-08-349-867-26

; Sequence 26, Application US/08349867  
; Patent No. 5508264  
; GENERAL INFORMATION:  
; APPLICANT: Bradfisch, Gregory A.  
; APPLICANT: Thompson, Mark  
; APPLICANT: Schwab, George E.  
; TITLE OF INVENTION: No. 5508264el Pesticidal Compositions  
; NUMBER OF SEQUENCES: 34  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: David R. Saliwanchik  
; STREET: 2421 N.W. 41st Street, Suite A-1  
; CITY: Gainesville  
; STATE: FL  
; COUNTRY: USA  
; ZIP: 32606  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/349,867  
FILING DATE:  
CLASSIFICATION: 514  
ATTORNEY/AGENT INFORMATION:  
NAME: Saliwanchik, David R.  
REGISTRATION NUMBER: 31,794  
REFERENCE/DOCKET NUMBER: MA86  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 904-375-8100  
TELEFAX: 904-372-5800  
INFORMATION FOR SEQ ID NO: 26:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 3444 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
US-08-349-867-26

Query Match 3.9%; Score 85.8; DB 2; Length 3444;

Best Local Similarity 53.8%; Pred. No. 5.1e-08;  
Matches 220; Conservative 0; Mismatches 162; Indels 27; Gaps 1;

QY 589 ATACGATTGGAGAAATGTTACAAATGATTTATTCGAGAAATACCTGGTTCCAACTTGA 648  
DB 388 ATTCGATTGGCTAATACAGACAGCGCTTAATACGCAATTAATTTACACTTCA 447  
QY 649 ACTTATAAACGCTATTACTTACTTATTAATGCGCAAGCTGCTAATTTCAATTAATTA 708  
DB 448 AGTTTGAATCCCTCTTTATCGGCTATGTTCAAGCGCGCAATTTACATTTACACTA 507  
QY 709 TTACAAACAAGTGCTGAATTTGCTGATGAAATGCAATATACATCTTCACTCAAA 768  
DB 508 TTAAGACGCGTATGCTGTTGGGAGGTTGGGAGCTGATATGCTATGTAATAT 567  
QY 769 GAACCTAATGCTGAACATCAGATGACTTATTAATCTTTAAAGAAATATACCTAA 828  
DB 568 C-----ATTATATAGATTATTAATCTTTATTTATGA 600  
QY 829 TATAGTAATCTATTTGCAAAATACCTATAGAACAGACTTAAATCTTAGAGACGAACA 888  
DB 601 TATAGCAAACTTTGTTGGACATACATCAAGATTAGAAACTTAGAAGCTACTAT 660  
QY 889 AATATGAATGAGATATTTATGACTATGCAAGATATATGACCACTTACTATTAAT 948  
DB 661 ACTGACAAATGGGCAAGATTCAATCAGTTAGAGAGATTAACTTACTGTTATGAT 720  
QY 949 ACCATCTCTCAATTTCTTTATATGATATTAATAAGATATAGATTCA 997  
DB 721 ATCGTGTCTTTTCCGAACTACATGTTAGAACTATCCAAATTCAAA 769

RESULT 9  
US-08-239-476-22

; Sequence 22, Application US/08239476  
; Patent No. 5527883  
; GENERAL INFORMATION:  
; APPLICANT: Thompson, Mark  
; APPLICANT: Schwab, George E.  
; TITLE OF INVENTION: Improvement of Delta-Endotoxin Expression in  
; NUMBER OF SEQUENCES: 34  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: David R. Saliwanchik  
; STREET: 2421 N.W. 41st Street, Suite A-1  
; CITY: Gainesville  
; STATE: Florida  
; COUNTRY: USA  
; ZIP: 32606

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/239,476  
FILING DATE:  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Saliwanchik, David R.  
REGISTRATION NUMBER: 31,794  
REFERENCE/DOCKET NUMBER: MA83  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (904) 375-8100  
TELEFAX: (904) 372-5800  
INFORMATION FOR SEQ ID NO: 22:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 3444 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
US-08-239-476-22

Query Match 3.9%; Score 85.8; DB 2; Length 3444;  
Best Local Similarity 53.8%; Pred. No. 5.1e-08;  
Matches 220; Conservative 0; Mismatches 162; Indels 27; Gaps 1;

QY 589 ATACGATTGGAAGTGTTCACATGATTTTATGAGAAATACCTGGTTCCACTGAA 648  
DB 388 ATTCGATTGCTAATACAGACGACGCTTTAATACGCAATTAATTTTACACTTCA 447  
QY 649 ACTATAAAGCTATTAATCTATTTATGCGAAGCTGCTAATTTTCAATTAAATTA 708  
DB 448 AGTTTGAATCCCTCTTTATCGGTATGTTCAAGCGGGAATTTACATTATACCTA 507  
QY 709 TTACAACAGGCTGTAATGGCTGATGAATGCAATATACATCTTCACAAAAT 768  
DB 508 TTAAGAGACGCTGATCCCTTTGCGAGGCTTGAGATATAGCTACTGTTAATAT 567  
QY 769 GAACCTAATGCTGGAACATCAGATGACTATTAATTAACCTTTAAAGAAATATACCTAAA 828  
DB 568 C-----ATTATATATAGATTAATTAATTAATCTTATTCATAGA 600  
QY 829 TATAGTAATATTGTGCAATACCTATAGAACAGACTAATAAAATCTTAGAGAGAACCA 888  
DB 601 TATAGCAAACTTGTGACACATACATCAAGATTTAGAAAATTAAGAGTACTAAT 660  
QY 889 AATATGAATGAGATATTTAATGACTATGAGAGATATATGACCAATTAATTTAGAT 948  
DB 661 ACTGACAAATGGGCAAGATCAATCAGTTAGAGAGATTTAACAACCTTACTGTTATGAT 720  
QY 949 ACCATCTCTCAATTTCTTTATATGATATTAAGATATAGATTTCAA 997  
DB 721 ATCGTGCTCTTTTCCGAACACTAGATGTAGAACATATCCAATTTCAA 769

RESULT 10  
US-08-239-476-26  
Sequence 26, Application US/08239476  
Patent No. 5527883  
GENERAL INFORMATION:  
APPLICANT: Thompson, Mark  
APPLICANT: Schwab, George E.  
TITLE OF INVENTION: Improvement of Delta-Endotoxin Expression in  
TITLE OF INVENTION: Pseudomonas fluorescens  
NUMBER OF SEQUENCES: 34  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: David R. Saliwanchik  
STREET: 2421 N.W. 41st Street, Suite A-1  
CITY: Gainesville  
STATE: Florida

COUNTRY: USA  
ZIP: 32606  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/239,476  
FILING DATE:  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Saliwanchik, David R.  
REGISTRATION NUMBER: 31,794  
REFERENCE/DOCKET NUMBER: MA83  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (904) 375-8100  
TELEFAX: (904) 372-5800  
INFORMATION FOR SEQ ID NO: 26:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 3444 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
US-08-239-476-26

Query Match 3.9%; Score 85.8; DB 2; Length 3444;  
Best Local Similarity 53.8%; Pred. No. 5.1e-08;  
Matches 220; Conservative 0; Mismatches 162; Indels 27; Gaps 1;

QY 589 ATACGATTGGAAGTGTTCACATGATTTTATGAGAAATACCTGGTTCCACTGAA 648  
DB 388 ATTCGATTGCTAATACAGACGACGCTTTAATACGCAATTAATTTTACACTTCA 447  
QY 649 ACTATAAAGCTATTAATCTATTTATGCGAAGCTGCTAATTTTCAATTAAATTA 708  
DB 448 AGTTTGAATCCCTCTTTATCGGTATGTTCAAGCGGGAATTTACATTATACCTA 507  
QY 709 TTACAACAGGCTGTAATGGCTGATGAATGCAATATACATCTTCACAAAAT 768  
DB 508 TTAAGAGACGCTGATCCCTTTGCGAGGCTTGAGATATAGCTACTGTTAATAT 567  
QY 769 GAACCTAATGCTGGAACATCAGATGACTATTAATTAACCTTTAAAGAAATATACCTAAA 828  
DB 568 C-----ATTATATATAGATTAATTAATTAATCTTATTCATAGA 600  
QY 829 TATAGTAATATTGTGCAATACCTATAGAACAGACTAATAAAATCTTAGAGAGAACCA 888  
DB 601 TATAGCAAACTTGTGACACATACATCAAGATTTAGAAAATTAAGAGTACTAAT 660  
QY 889 AATATGAATGAGATATTTAATGACTATGAGAGATATATGACCAATTAATTTAGAT 948  
DB 661 ACTGACAAATGGGCAAGATCAATCAGTTAGAGAGATTTAACAACCTTACTGTTATGAT 720  
QY 949 ACCATCTCTCAATTTCTTTATATGATATTAAGATATAGATTTCAA 997  
DB 721 ATCGTGCTCTTTTCCGAACACTAGATGTAGAACATATCCAATTTCAA 769

RESULT 11  
US-08-598-305A-22  
Sequence 22, Application US/08598305A  
Patent No. 5827514  
GENERAL INFORMATION:  
APPLICANT: BRADFICH, Gregory A.  
APPLICANT: THOMPSON, Mark  
APPLICANT: SCHWAB, George E.  
TITLE OF INVENTION: No. 5827514el Pesticidal Compositions  
NUMBER OF SEQUENCES: 38  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Saliwanchik, Lloyd & Saliwanchik  
STREET: 2421 N.W. 41st Street, Suite A-1

CITY: Gainesville  
STATE: FL  
COUNTRY: USA  
ZIP: 32606-6669  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/598,305A  
FILING DATE: 08-FEB-1996  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/349,867  
FILING DATE: 06-DEC-1994  
CLASSIFICATION: 514  
ATTORNEY/AGENT INFORMATION:  
NAME: Sanders, Jay M.  
REGISTRATION NUMBER: 39,355  
REFERENCE/DOCKET NUMBER: MA86.D1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 352-375-8100  
TELEFAX: 352-372-5800  
INFORMATION FOR SEQ ID NO: 22:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 3444 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
US-08-598-305A-22

Query Match 3.9%; Score 85.8; DB 2; Length 3444;  
Best Local Similarity 53.8%; Pred. No. 5.1e-08;  
Matches 220; Conservative 0; Mismatches 162; Indels 27; Gaps 1;  
QY 589 ATACGATTGGAGAAATGTTACAAATGATTTTATTCGAGAAATACCTGGTTTCCAACTGAA 648  
DB 388 ATTGCATTGGCTAATACAGACAGACGCTTTAATACAGCAATTAATTTTACACTTACA 447  
QY 649 ACTTATAAAGCGCTAATTAATCTAATTTATGCGCAAGCTGCTAATTTTCATTAAATTTA 708  
DB 448 AGTTTGAATCCCTCTTTATCGGCTATGTTCAAGCGGCGAATTTACATTATCACCTA 507  
QY 709 TTACACAAGGTCGTAATTTGCTGATGAATGCAATATACATCCCTTCACAAT 768  
DB 508 TTAAGAGAGCGCTGATCGTTGGGAGGTTGGGAGCTGATATGCTACTGTTATATAT 567  
QY 769 GAACCTAATGCTGGAACATCAGATGACTATTATTAACCTTTAAAGAAATATACCTAA 828  
DB 568 C-----ATTATATATAGATTATATTAATCTTATTTATGA 600  
QY 829 TATAGTAATCTATTTGCAAAATACCTATAGAACAGACTAAAAATCTTAGAGACGAACA 888  
DB 601 TATAGCAAAATGTTGTTGAGACATACATCAAGATTAGAAAACCTTAAGAGGTAAT 660  
QY 889 AATATGAATGAGATATTTTATATGACTATGCAAGATATATGACCATTAAGTATTAAGT 948  
DB 661 ACTGACAAATGGGCAAGATTCATCAGTTAGGAGAGATTTAAACCTTAAGTATTAAGT 720  
QY 949 ACCATCTCAATTTTCTTTATATGATATTAAGATTTAGAGATTCA 997  
DB 721 ATCGTGTCTTTTTCGAACTACGATGTAGAACATATCCATTCAAA 769

RESULT 12  
US-08-598-305A-26  
Sequence 26, Application US/08598305A  
Patent No. 5827514  
GENERAL INFORMATION:  
APPLICANT: BRADFISCH, Gregory A.  
APPLICANT: THOMPSON, Mark

APPLICANT: SCHWAB, George E.  
TITLE OF INVENTION: No. 5827514e1 Pesticidal Compositions  
NUMBER OF SEQUENCES: 38  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Saliwanchik, Lloyd & Saliwanchik  
STREET: 2421 N.W. 41st Street, Suite A-1  
CITY: Gainesville  
STATE: FL  
COUNTRY: USA  
ZIP: 32606-6669  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/598,305A  
FILING DATE: 08-FEB-1996  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/349,867  
FILING DATE: 06-DEC-1994  
CLASSIFICATION: 514  
ATTORNEY/AGENT INFORMATION:  
NAME: Sanders, Jay M.  
REGISTRATION NUMBER: 39,355  
REFERENCE/DOCKET NUMBER: MA86.D1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 352-375-8100  
TELEFAX: 352-372-5800  
INFORMATION FOR SEQ ID NO: 26:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 3444 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
US-08-598-305A-26

Query Match 3.9%; Score 85.8; DB 2; Length 3444;  
Best Local Similarity 53.8%; Pred. No. 5.1e-08;  
Matches 220; Conservative 0; Mismatches 162; Indels 27; Gaps 1;  
QY 589 ATACGATTGGAGAAATGTTACAAATGATTTTATTCGAGAAATACCTGGTTTCCAACTGAA 648  
DB 388 ATTGCATTGGCTAATACAGACAGACGCTTTAATACAGCAATTAATTTTACACTTACA 447  
QY 649 ACTTATAAAGCGCTAATTAATCTAATTTATGCGCAAGCTGCTAATTTTCATTAAATTTA 708  
DB 448 AGTTTGAATCCCTCTTTATCGGCTATGTTCAAGCGGCGAATTTACATTATCACCTA 507  
QY 709 TTACACAAGGTCGTAATTTGCTGATGAATGCAATATACATCCCTTCACAAT 768  
DB 508 TTAAGAGAGCGCTGATCGTTGGGAGGTTGGGAGCTGATATGCTACTGTTATATAT 567  
QY 769 GAACCTAATGCTGGAACATCAGATGACTATTATTAACCTTTAAAGAAATATACCTAA 828  
DB 568 C-----ATTATATATAGATTATATTAATCTTATTTATGA 600  
QY 829 TATAGTAATCTATTTGCAAAATACCTATAGAACAGACTAAAAATCTTAGAGACGAACA 888  
DB 601 TATAGCAAAATGTTGTTGAGACATACATCAAGATTAGAAAACCTTAAGAGGTAAT 660  
QY 889 AATATGAATGAGATATTTTATATGACTATGCAAGATATATGACCATTAAGTATTAAGT 948  
DB 661 ACTGACAAATGGGCAAGATTCATCAGTTAGGAGAGATTTAAACCTTAAGTATTAAGT 720  
QY 949 ACCATCTCAATTTTCTTTATATGATATTAAGATTTAGAGATTCA 997  
DB 721 ATCGTGTCTTTTTCGAACTACGATGTAGAACATATCCATTCAAA 769

RESULT 13

```
US-08-639-923A-22
; Sequence 22, Application US/08639923A
; Patent No. 5840554
; GENERAL INFORMATION:
; APPLICANT: Thompson, Mark
; APPLICANT: Schwab, George B.
; TITLE OF INVENTION: Improvement of Delta-Endotoxin Expression in
; TITLE OF INVENTION: Pseudomonas fluorescens
; NUMBER OF SEQUENCES: 38
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: David R. Saliwanchik
; STREET: 2421 N.W. 41st Street, Suite A-1
; CITY: Gainesville
; STATE: Florida
; COUNTRY: USA
; ZIP: 32606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/639,923A
; FILING DATE: 24-APR-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/239,476
; FILING DATE: 06-MAY-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Saliwanchik, David R.
; REGISTRATION NUMBER: 31,794
; REFERENCE/DOCKET NUMBER: M83.D1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (904) 375-8100
; TELEFAX: (904) 372-5800
; INFORMATION FOR SEQ ID NO: 22:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3444 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; US-08-639-923A-22

Query Match          3.9%; Score 85.8; DB 2; Length 3444;
Best Local Similarity 53.8%; Pred. No. 5.1e-08;
Matches 220; Conservative 0; Mismatches 162; Indels 27; Gaps 1;
```

```
QY 949 ACCATCTGCATTTCTTTATATGATATATAAAGATATAGAGTTCAA 997
DB 721 ATCGTGTCTTTTCCGAAGTATGAGATATGACATATCAATTCAA 769

RESULT 14
US-08-639-923A-26
; Sequence 26, Application US/08639923A
; Patent No. 5840554
; GENERAL INFORMATION:
; APPLICANT: Thompson, Mark
; APPLICANT: Schwab, George B.
; TITLE OF INVENTION: Improvement of Delta-Endotoxin Expression in
; TITLE OF INVENTION: Pseudomonas fluorescens
; NUMBER OF SEQUENCES: 38
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: David R. Saliwanchik
; STREET: 2421 N.W. 41st Street, Suite A-1
; CITY: Gainesville
; STATE: Florida
; COUNTRY: USA
; ZIP: 32606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/639,923A
; FILING DATE: 24-APR-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/239,476
; FILING DATE: 06-MAY-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Saliwanchik, David R.
; REGISTRATION NUMBER: 31,794
; REFERENCE/DOCKET NUMBER: M83.D1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (904) 375-8100
; TELEFAX: (904) 372-5800
; INFORMATION FOR SEQ ID NO: 26:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3444 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; US-08-639-923A-26

Query Match          3.9%; Score 85.8; DB 2; Length 3444;
Best Local Similarity 53.8%; Pred. No. 5.1e-08;
Matches 220; Conservative 0; Mismatches 162; Indels 27; Gaps 1;
```



Db 601 TATACGAACATGTTGGACATACATCAAGATTGAAAACCTTAAGAGTACTAAT 660  
Qy 889 AATATGAATGAGATATATTTATGACTATCGAAGATATATGACCATTCGTATTTAGT 948  
Db 661 ACTGACAAATGGGCAAGATTCATCACTTTAGAGAGATTTAACCTTACTGTATTTAGT 720  
Qy 949 ACCATCTCTCAATTTCTTTATATGATATATATAAAGATATAGAGATTCAA 997  
Db 721 ATCGTTGCTCTTTTCCGAACATGATGTAGAACATATCCAAATTCAAA 769

RESULT 15  
PCT-US95-05431-22  
; Sequence 22, Application PC/TUS9505431  
; GENERAL INFORMATION:  
; APPLICANT:  
; APPLICANT: Street address: 5501 Oberlin Drive  
; APPLICANT: City: San Diego  
; APPLICANT: State/Province: California  
; APPLICANT: Country: US  
; APPLICANT: Postal code/Zip: 92121  
; APPLICANT: Phone number: (619) 453-8030 Fax number: (619)453-6991  
; APPLICANT: Telex number:  
; TITLE OF INVENTION: Improvement of Delta-Endotoxin Expression in  
; TITLE OF INVENTION: Pseudomonas fluorescens  
; NUMBER OF SEQUENCES: 34  
; CORRESPONDENCE ADDRESS:  
; ADDRESSER: David R. Saliwanchik  
; STREET: 2421 N.W. 41st Street, Suite A-1  
; CITY: Gainesville  
; STATE: Florida  
; COUNTRY: USA  
; ZIP: 32606  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US95/05431  
; FILING DATE:  
; CLASSIFICATION:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Saliwanchik, David R.  
; REGISTRATION NUMBER: 31,794  
; REFERENCE/DOCKET NUMBER: MA83  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (904) 375-8100  
; TELEFAX: (904) 372-5800  
; INFORMATION FOR SEQ ID NO: 22:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 3444 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: DNA (genomic)  
PCT-US95-05431-22

Query Match 3.9%; Score 85.8; DB 6; Length 3444;  
Best Local Similarity 53.8%; Pred. No. 5.1e-08;  
Matches 220; Conservative 0; Mismatches 162; Indels 27; Gaps 1;

Qy 589 ATACGATTTGAGAAATGTTCAAAATGATTTTATTCGAAATACCTGGTTCCAACTTGAA 648  
Db 388 ATTCGATTTGCTAATACAGACAGCGCTTAATACAGCAATTAATTAATTTACACTTACA 447  
Qy 649 ACTATATAAACGCTATTACTATTTATGCGCAAGCTGTAATTTTCAATTTAAATTTA 708  
Db 448 AGTTTGAATATCCCTCTTTATCGGTCTATGTTCAAGCGGCAATTTACATTTATCACATA 507  
Qy 709 TTACAACAAGGTGCTGAATGGCTGATGAATGAATGCAATATACATCTTCACAAATT 768  
Db 508 TTAAGAGCGGTGATCGTTGGGCGAGGGTGGGCACTGATATATAGTACTGTATAAAT 567

Qy 769 GAACCTAATGCTGGAACATCAGATGACTATTATTAATCTTTAAAGAAAATATACCTAAA 828  
Db 568 C-----ATTATATATAGATTATATTAATCTTATTTACATA 600  
Qy 829 TATAGTAATCTATTTGCAAAATACCTATAGAAACAGGACTPAAAAATCTTAGACAGAACCA 888  
Db 601 TATACGAACAATTTGTTGGACACATACATCAAGGATTAGAAAACCTTAAGAGGTACTAAT 660  
Qy 889 AATATGAATGAGATATATTTATATGACTATGCAAGATATATAGCACTACTGTATTTAGAT 948  
Db 661 ACTGACAAATGGGCAAGATTCATCACTTTAGAGAGATTTAACACTTACTGTATTTAGAT 720  
Qy 949 ACCATCTCTCAATTTCTTTATATGATATATAAAGATATAGAGATTCAA 997  
Db 721 ATCGTTGCTCTTTTCCGAACATGATGTAGAACATATCCAAATTCAAA 769

Search completed: December 11, 2005, 10:14:37  
Job time : 411 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 11, 2005, 13:37:06 / Search time 8805 Seconds  
(without alignments)  
11732.638 Million cell updates/sec

Title: US-10-783-417-1  
Perfect score: 2208  
Sequence: 1 atgatacaataacagataa.....atcatcaaaacacttga 2208

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 41078325 seqs, 23393541228 residues

Total number of hits satisfying chosen parameters: 82156650

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

EST:\*  
1: gb\_est1:\*  
2: gb\_est2:\*  
3: gb\_est3:\*  
4: gb\_hcc:\*  
5: gb\_est4:\*  
6: gb\_est5:\*  
7: gb\_est6:\*  
8: gb\_est7:\*  
9: gb\_gest1:\*  
10: gb\_gest2:\*  
11: gb\_gest3:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	94.6	4.3	1392	10	CG757503 P052-4-C0
C 2	90	4.1	1254	10	AG349719 Mus muscu
C 3	88.2	4.0	1758	10	CU509408 SAIL 811
C 4	87.6	4.0	1896	10	CG753083 P048-1-C0
C 5	87.2	3.9	1388	10	AG278124 Mus muscu
C 6	86.8	3.9	1101	10	CNS00396
C 7	84.6	3.8	1348	10	CG749499
C 8	84	3.8	1101	10	CNS006VL
C 9	82.8	3.7	1238	10	AJ925855
C 10	82.4	3.7	1489	10	AG350139 Mus muscu
C 11	82.4	3.7	1594	10	CU110653
C 12	82.2	3.7	1277	9	CC253231
C 13	80.4	3.6	1268	10	AG347098 Mus muscu
C 14	80	3.6	1242	10	CU068807
C 15	79.6	3.6	956	10	CG938578 TC818.3 G
C 16	79.4	3.6	1101	10	AL063921 Drosophila
C 17	78.4	3.6	2157	10	CU081966 CH216-165
C 18	78.2	3.5	1391	10	CG754863 P050-2-C0
C 19	77	3.5	1592	10	CG750135 P044-3-D0
C 20	76.8	3.5	1581	10	CU082000 CH216-165
C 21	76.6	3.5	994	11	CNS04N0J
C 22	76.6	3.5	1325	10	CG952344 TC838.1_D

C 23	76	3.4	1193	10	CG745316 P038-1-G0
C 24	75.6	3.4	1189	10	AJ925925
C 25	75.6	3.4	1227	10	AG430010 Mus muscu
C 26	74.6	3.4	1352	10	AG381852 Mus muscu
C 27	74.4	3.4	734	10	CNS010MP
C 28	74.2	3.4	1101	10	CG942516
C 29	74.2	3.4	1359	8	DN685273
C 30	74	3.4	1052	10	AJ927993
C 31	73.8	3.3	1542	10	AG386981 Mus muscu
C 32	73.6	3.3	1594	10	CU084406
C 33	73.4	3.3	1076	1	AJ926041
C 34	73.2	3.3	1190	10	CNS020N7
C 35	73.2	3.3	1696	10	AG346840 Mus muscu
C 36	73.2	3.3	1928	10	CU073845
C 37	73	3.3	1101	10	CNS00K85
C 38	73	3.3	1378	8	DN736559 CNB87-E11
C 39	72.8	3.3	1083	1	AJ926215
C 40	72.8	3.3	1260	10	CU491610
C 41	72.6	3.3	1101	10	CNS006VL
C 42	72.4	3.3	1539	10	AG340947 Mus muscu
C 43	72.2	3.3	1101	10	CNS017V2
C 44	72.2	3.3	1140	10	CU071552
C 45	72.2	3.3	1378	10	AG350209 Mus muscu

#### ALIGNMENTS

RESULT 1  
CG757503/c 1392 bp DNA linear GSS 24-OCT-2003  
LOCUS P052-4-C08.za Ppa EcORI BAC library Pristionchus pacificus genomic,  
DEFINITION genomic survey sequence.  
ACCESSION CG757503  
VERSION CG757503.1 GI:37986131  
KEYWORDS GSS.  
SOURCE Pristionchus pacificus  
ORGANISM Pristionchus pacificus  
Eukaryota; Metazoa; Nematoda; Chromadorea; Diplogasterida;  
Neodiplogasteridae; Pristionchus.

#### REFERENCE

1 (bases 1 to 1392)  
Srinivasan,J., Sinz,W., Jeeze,T., Wiggers-Perebolle,L., Jansen,K.,  
Buntjer,J., Van der Meulen,M. and Sommer,R.J.  
An integrated physical and genetic map of the nematode Pristionchus

#### TITLE

pacificus  
Mol. Genet. Genomics 269 (5), 715-722 (2003)

#### JOURNAL

PUBMED

COMMENT  
Contact: Sommer RJ  
Evolutionary Biology  
Max-Planck-Institute for Developmental Biology  
Spemannstr. 37-39, Tuebingen D-72076, Germany  
Tel: 00497071601371  
Fax: 00497071601498  
Email: ralf.sommer@tuebingen.mpg.de

#### FEATURES

source  
Location/Qualifiers  
1..1392  
/organism="Pristionchus pacificus"  
/mol\_type="genomic DNA"  
/strain="California"  
/db\_xref="taxon:54126"  
/clone\_lib="Ppa EcORI BAC library"  
/note="The library was generated by a partial digest of  
the genomic DNA with EcORI and cloning into the BAC  
vector."

#### ORIGIN

Query Match 4.3% Score 94.6; DB 10; Length 1392;  
Best Local Similarity 44.6%; Pred. No. 7.5e-08;  
Matches 536; Conservative 0; Mismatches 654; Indels 12; Gaps 4;

360 AGAACAGACAAACAGATGAGACAAATTTATTAATGGAGAAATTTTGTTCATGC 419

[illegible]

Qy	191	TAGAAAAAGACGCAATCTAGTATTGAATCCAGGTGGTCACCAACTTAATACTA	1550
Dd	230	AAAAAAAAAAATATATTTATTTTAAATTTAAAATTTAAAAAAATTTAAATAATA	171
Qy	1551	TA 1552	
Dd	170	TA 169	

RESULT 2					
AG349719/c					
LOCUS					
DEFINITION	AG349719	1254 bp	DNA	linear	GSS 18-DEC-2004
	Mus musculus molossinus				
	sequence.				
ACCESSION	AG349719				
VERSION	AG349719.1	GI:47923029			
KEYWORDS	GSS.				
SOURCE	Mus musculus molossinus (Japanese wild mouse)				

## REFERENCE

**TITLE** Contribution of Asian mouse subspecies *Mus musculus molossinus* to

genomic constitution of strain C57BL/6J, as defined by BAC-end  
Sequence-SNP analysis  
Genome Res. 14 (12), 2439-2447 (2004)

PUBMED	15574823
REFERENCE	2 (bases 1 to 1254)
AUTHORS	Hattori, M., Toyoda, A., Noguchi, H., Kojima, T. and Sakaki, Y

**TITLE**  
**JOURNAL**

and Chemical Research (RIKEN), Genomic Sciences Center (GSC), 1-7-22 Suehiro-chou, Teikuri-ku, Yokohama, Kanagawa, 230-0045, Japan (E-mail: hatori@gs.c.riken.jp, URL: <http://ngp.gs.c.riken.go.jp/>, Tel: 81-45-503-9111, Fax: 81-45-503-9170)  
Clones are derived from the mouse BAC library M5ng01. For BAC

library availability, please contact Kuniya Abe (abe@rtc.riken.jp)  
Tsukuba Institute, Bio Resource Center,  
The Institute of Physical and Chemical Research (RIKEN) 3-1-1  
Koyadai, Tsukuba, 305-0074 Japan  
phone: 81-298-36-9189, fax: 81-298-36-9199  
e-mail: abe@rtc.riken.jp

PRIMERS  
Sequencing : T7  
LIBRARY

Vector	: pBACE3.6
R.Site 1	: EcoRI
R.Site 2	: EcoRI.
FEATURES	Location/Qualifiers

```
source
1. .1254
/organism="Mus musculus molossinus"
/mol_type="genomic DNA"
```

```

/sub_species="molossinus"
/db_xref="taxon:57486"
/clone="MSMg01-146B10.T7"
/sex="male"
/tissue_type="mixture of kidney and spleen"
/clone_lib="MSMg01 Mouse male BAC library"

```

## ORIGIN

Query Match 4.1% Score 90 DB 10 Length 1254

Best Local Similarity	43.28;	Pred. No. 5.8e-07;	
Matches 480;	Conformations 0;	Mismatches 530;	Indels 3;
			Gaps 11;

	Matches	100%	Conserved	% mismatches	90%	Inserts	2%	Gaps	1%
Q7	391	ATTAAATGGAGAAATTTTGTGTAACACCCTTAACAAGAACATAAACAGCTAAAG	450						
bB	1352	ATAAATATATATAAATAATTAATAATTTAAATTAATATCTAATTAATAATATACAAAT	1193						

451 TTACAACTTTAGAGATTAGACAAATATTACAAAGCTATTAACAGATTAGATGAT 510  
 1192 AATTAAATATATAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1133  
 511 TGGAGAAATTTAAAGACTACAGACTCTGATTAACCACTCATCTGAGATTACACA 570  
 1132 TTAATAATATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1073  
 571 GCTGCTTGACCTTAATAATACGATTTGGAAATGTTCACAATGATTTTATTCGAGAA 630  
 1072 ATATTAAATTAATAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1013  
 631 CCTGCTTCCAACTTGAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 690  
 1012 AATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 953  
 631 AATTTCATTTAATTTATTAACACAGAGTGTGATTTGGCTGATGATGATGATGAT 750  
 952 AATTAATAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 893  
 751 ATACATCCTTGACAAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 810  
 892 AAAAAAAAAATTAATAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 833  
 811 AAGAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 870  
 832 TAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 773  
 871 AATCTTAGACGAACTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 930  
 772 AAAAAAAAAATTAATAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 713  
 931 AACTTACTGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 990  
 712 AAAAAAAAAATTAATAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 653  
 991 GATTCATTAAGAGAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1050  
 652 AATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 593  
 1051 ACAAATGAATTAATTTGATGCTCTCTCACTTAAGATTCACCAATCTAGCTACG 1110  
 592 AATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 533  
 1111 ATGAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1170  
 532 TAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 473  
 1171 TTTTATTAAGAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1230  
 472 TATATATACACAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 413  
 1231 CTTACTTAAGCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1290  
 412 TTTTCTCTAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 353  
 1291 ACAAATAAGAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1348  
 352 AATTTTACATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 293  
 1349 CACCTCTGTTCCCTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1408  
 292 ATTAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 233  
 1409 ATGCTCATCTTAACACACACTCAAAATTAATTAATTAATTAATTAATTAATTAATTA 1468  
 232 ATATTAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 173  
 1469 ACAACAATTTTCTTAATTTCTAGAAAAAA 1500  
 172 ATTCAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 141

RESULT 3  
 CL509408  
 LOCUS  
 DEFINITION  
 SAIL\_811\_H11.v3 SAIL Collection Arabidopsis thaliana genomic clone  
 SAIL\_811\_H11.v3, genomic survey sequence.  
 CL509408  
 CL509408.1 GI:4606728  
 GSS.  
 SOURCE  
 ORGANISM  
 Arabidopsis thaliana (thale cress)  
 Arabidopsis thaliana  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;  
 rosoids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.  
 REFERENCE  
 1 (bases 1 to 1758)  
 Sessions A., Burke E., Presting G., Aux G., McIlver J., Paton D.,  
 Dietrich B., Ho P., Bawaden J., Ko C., Clarke J.D., Cotton D.,  
 Bullis D., Snell J., Mignel T., Hutchinson D., Kimerly B.,  
 Mitzel T., Katagiri F., Glazebrook J., Law M. and Goff S.A.  
 A high-throughput Arabidopsis reverse genetics system  
 Plant Cell 14 (12), 2985-2994 (2002)  
 12468722  
 JOURNAL  
 PUBMED  
 COMMENT  
 Contact: Sessions A  
 Applied Trait Genetics  
 Syngenta Biotechnology Inc.  
 3054 Cornwallis Rd., Research Triangle Park, NC 27709, USA  
 Email: allen.sessions@syngenta.com  
 ABRC Stock Number CS836276; T-DNA left border flanking sequences of  
 Syngenta Arabidopsis Insertion Library (SAIL) lines are available  
 through the Arabidopsis Biological Resource Center (ABRC).  
 Sequences represent a pool of amplified genomic regions and not  
 single contiguous sequences.  
 Class: TDNA tagged.  
 FEATURES  
 source  
 location/Qualifiers  
 1..1758  
 /organism="Arabidopsis thaliana"  
 /mol\_type="genomic DNA"  
 /ecotype="Columbia"  
 /db\_xref="taxon:3702"  
 /clone="SAIL\_811\_H11.v3"  
 /clone\_lib="SAIL\_Collection"  
 /note="T-DNA left border sequences were isolated using a  
 modified Tail-PCR strategy"  
 ORIGIN  
 Query Match 4.0%; Score 88.2; DB 10; Length 1758;  
 Best Local Similarity 33.0%; Pred. No. 1.3e-06;  
 Matches 452; Conservative 0; Mismatches 908; Indels 10; Gaps 1;  
 382 ACACAAATTTAATAATGAGAAATTTGTTGATACACCGTTTACAGAAAGCATATAA 441  
 131 AA 190  
 442 CAGCTAAAGTTACAACTTTAGAAAGATTAGACAAATATTACAAAGCTATTAATACACA 501  
 191 NNN 250  
 502 TTAAGATGTTGAGAGAAATTTAAAGACTACAGCTCTGATTAACACATCAACACACA 561  
 251 AA 310  
 562 TTACAACAGCTGCTTGACCTTAATAATACGATTTGGAATGTTCAACATGATTTTAT 621  
 311 AATAATATATATTTT 370  
 622 CGAGAAATACCTGTTTCCAACTTGAAACTTAATAAAGCTATTAATTAATTAATGCG 681  
 371 TATTAATTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 430  
 682 CAAGCTGTAATTTTCAATTAATTAATTAACACAGAGTGTGATGATGATGATGATG 741  
 431 AAATAAATAAATAA 490  
 742 AATGAGATATATACCTTCAAAATTAATTAATGCTGGAACATGATGATATATAT 801

D	b		491	AAAAAAAAAANNAAAAAANNNNNAANAAAAANTTAAANAANNAANNAANNTTAAN	550
Q	y		802	AAACTTTTAAAAGAAATATACTTAATATAGTAACTATTGCAATPACTATGAACA	861
D	b		551	AAAAAAAAAANNNAAAAANNNNNAANNAANNNNNAAAAANNAANNAANNNNNN	610
Q	y		862	GGACTTAAAAAATCTTAGAGACGAACCATAATGAAATGAGATATTTTAAGACTAGCA	921
D	b		611	NN	670
Q	y		922	AGATATATGACCATTACTGTATATAGATCCATCTCTCAATTTCTTTATATGATTA	981
D	b		671	NNNNNAANNNNNAAAAANNNNNNAANAAAAANNNNNNNNNNNNNNNNNNN	730
Q	y		982	AGATTATAGATTCATAGAGAATAGAAATGTAAGCATTAAGATGAAGCTCACA	1041
D	b		731	AAAAAAAAAANNNNNAAAAANNNNNNAANNAANNAANNAANNNNNNNNN	790
Q	y		1042	GAAATTTTATACACTGAATTAATTTTGATCGTCTCCCTCACTTAGAGTTCA	1107
D	b		791	NNNNAAANNNNTTNNAATAAANNNNAANNAATTAATAATTAATTAATTAAT	850
Q	y		1102	CTAGCTAGCATGGAATATATTTAACACGTGCAAGTTTAAATTAATTTTCATTT	1166
D	b		851	ATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT	910
Q	y		1161	-----ACAAATTAATTTTTTATATACAGAAAATCAATTTGGGAATCGTTAGTGG	1211
D	b		911	TNNNNAAATTAATTAANAAAAATTTAATTAATTAATTAATTAATTAATTAAT	970
Q	y		1212	TATTTCTAATCGTATGACCTACTTAATAGCAATATCTAACTGAACTTTATAT	1271
D	b		971	TTAATTAATTAANNTTAATTAATTAATTAATTAATTAATTAATTAATTAAT	1030
Q	y		1272	AAGAAGAGGTCACCCACAAACAAAACATAAGACATTTGAATCTTAAGTTCA	1333
D	b		1031	NNNNAAANNTATTAATTAANNTATATATTAATTAATTAATTAATTAATTAAT	1090
Q	y		1332	TGTAATGATAGACATCACCTCGTGTCCCTATTCACACACACTTAATTAATTA	1391
D	b		1091	NATTAATTAANNTTTTAATTAATTAATTAATTAATTAATTAATTAATTAAT	1150
Q	y		1392	AATTGAATTTATTTAAATGGCTACTTAACAACACACTCAAATTTTCAGCAG	1451
D	b		1151	TAAATTAANNAATTAATTAANNTTAATTAATTAATTAATTAATTAATTAAT	1210
Q	y		1452	TTTATCTAATATCAAAAACACACTTTTTTCAATTTCCTAGAAAAAAGACTG	1511
D	b		1211	AATATNTTAANTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT	1270
Q	y		1512	AGTTATGATCGAGGTGGTCCAAACTTTAATATATATAGTCAATTTATCCAT	1571
D	b		1271	AANATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT	1330
Q	y		1572	TTCAATTAATTAATTAATTCCTAATGCTATGGAATTAAGCTAACAATTAAT	1631
D	b		1331	ATTNNNTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT	1390
Q	y		1632	ATTAGATGACACACAGTAGTGTGATAGATATATATGCAATATCAGATTAAT	1691
D	b		1391	ATTNTTATTAATTAATTAATTTNNNTATATTAATTAATTAATTAATTAAT	1450
Q	y		1692	AATGATCCAGCATCAAAAGGTAAACATCTTGATCAAACTCTAAGGTAA	1741
D	b		1451	AATTAATTAANAAAAATTAATTAANAAAAATTAATTAATTAATTAATTAAT	1500

RESULT 4	CG753083/c	CG753083	1896 bp	DNA	linear	GSS 24-OCT-2003
LOCUS						
DEFINITION		P0481-C01.2a Ppa EcORI BAC Library <i>Pristionchus pacificus</i> genomic/c.				
		genomic survey sequence.				

ACCESSION	C6753083
VERSION	C6753083.1 GI:37977199
KEYWORDS	GSS.
SOURCE	Pristionchus pacificus Pristionchus pacificus <i>Eukaryota; Metazoa; Nematoda; Chromadorea; Diplogasterida; Neodiplostegidae;</i> Pristionchus. 1 (bases 1 to 1896) Srinivasan,J., Sins,W., Jesse,T., Wiggers-Perebolte,L., Jansen,K., Buntjer,J., van der Meulen,M. and Sommer,R.J. An integrated physical and genetic map of the nematode Pristionchus pacifcus
JOURNAL	MOL. Genet. Genomics 269 (5), 715-722 (2003)
PUBMED	12884007
COMMENT	Contact: Sommer RU Evolutionary Biology Max-Planck-Institute for Developmental Biology Spemannstr. 37-39, Tuebingen D-72076, Germany Tel.: 00497071601371 Faxl: 00497071601498 Email: ralf.sommer@tuebingen.mpg.de Class: BAC ends.
FEATURES	location/Qualifiers
source	1..1896 /organism="Pristionchus pacificus" /mol_type="genomic DNA" /strain="California" /db_xref="taxon:54126" (clone_1lb="Ppa EcoRI BAC library") /note="The library was generated by a partial digest of the genomic DNA with EcoRI and cloning into the BAC vector."
ORIGIN	
Query Match	4.0%; Score 87.6; DB 10; Length 1896;
Best Local Similarity	44.8%; Pred. No. 1.7e-06;
Matches 632:	Conservative 0; Mismatches 762; Indels 17; Gaps 9;
Dy	800 ATAACTTTTAAAAGAAAAATATCCTTAATATATGATACTATTGTGCMAATACCCTAATGA A 859
Dd	1821 AAAAATTTTTTAAATAAAAAT 1762
QY	860 CAGACTCAAAAATTCCTTAGAGAGAACAATATGAAATGGAGTA--TATTTAATGACTA 917
Dd	1761 AAATATATAATTAATATANAATTTAAATATAATATATATATATATATATATATATATATATIN 1702
QY	918 TCGAAGATATATGACCATTCGTGTAGATAGATACCACCTCCAATTTCTTATATATGATAT 977
Dd	1701 ATAATAATATATAAATTTAATAATATATATATATATATATATATATATATATATATATAT 1642
QY	978 AAAAAGATATAGAGATTCATATAGAGAAATAGAAAGTAAAGCATTPAAGATGAACAC 1037
Dd	1641 TAANNAAAAAAAAAATATATTTTAAATTTATATATATATATATATATATATATATATATAT 1582
QY	1038 AAGAGAAATTTATACAACTGAATTAATTTTGATTCGCTCCCACTTAGAGTTCAAC 1097
Dd	1581 TATATAATATATATATAT-TATTTTATATATTTTATTTTATATATATATATATATATTTTAA 1523
QY	1098 CAATCTAGCTAGATGAAATATATATTAACAAGCTGCAAGTTTAAATATATTCATTTT 1157
Dd	1522 TATTTTATTTAAAAAATATATAT-TAATATAATATATATTTTTTTTTTTTTTTTTAT 1465
QY	1158 AGAACATTTATTTTATATACAAAAATACAATTTCCGAATTCGTTAATGTTGATTC 1217
Dd	1464 NTAAAAATATATATTTTATATATAATTAATTAATTAATTAATTAATTAATTAATTAATTTTA 1405
QY	1218 TAAATCGATGACCTCTATATAGCATATCATPACTGAACCTTATATGAGAGAAAGAC 1277
Dd	1404 TTITATTAATATATTTTAAATATATATTTTAAATTTNANNAAATTAATATATATATATATAT 1345
QY	1278 AGGTGCCCAACAACAAAAACAATAGACCAATTTGAATCTTATPAAGTTCAAATGTGAC 1337
Dd	1344 ATATATTAATTAATTAATTAATATATTTTATTTTAAATTTTATTAATTAATTAATTTTAAATAA 1285

```

QY 1338 TGATAGACATCAGCTCTGTTCCCTATTCACACACTTAAATCAATGGA 1397
Db 1284 AAAAAAAAAATTTATATATATATATATATATATATATATATATAT 1225
QY 1338 ACTTATTTAAATGGCTCACTTAACACACTCAATTTGACAGAGGGCTTTATC 1457
Db 1224 ATATATTT--AATATATATATATATATATATATATATATATATATAT 1167
QY 1458 TAATATCAAAACAGACTTTTTCATTCATTCGTAAGAAAAGCTGAACTGAT 1517
Db 1166 TAATATATATATATATATATATATATATATATATATATATATATATAT 1107
QY 1518 TGATCAGGTTGTCACCAACTTAAATATATATATATATATATATATATAT 1577
Db 1106 AAAATATATATATATATATATATATATATATATATATATATATATATAT 1047
QY 1578 ATTTACTATTTCTATGAGATTGATTAACGCTAACAATATATATATATAT 1637
Db 1046 ATATTTTATATATATATATATATATATATATATATATATATATATAT 987
QY 1638 ATGACACACAGTATGTTGATATATATATATATATATATATATATATAT 1697
Db 966 ATATATATATATATATATATATATATATATATATATATATATATATAT 927
QY 1698 CCCAGCAATCAAAAGGTAACATCTTGATACAACTTAAAGTAATGGAAGC 1757
Db 926 AAAAAATATATATATATATATATATATATATATATATATATATATATAT 869
QY 1758 TACAGAGAGAACTTGTTATTTTCAAAAGTCAAGGGCTTTAGAAATTA 1817
Db 868 AAAAAATATATATATATATATATATATATATATATATATATATATATAT 809
QY 1818 TCCTATTTCTACACATCTTATATATATATATATATATATATATATAT 1877
Db 808 ATATATATATATATATATATATATATATATATATATATATATATATAT 753
QY 1878 TACTCTCTATATATATCTCTTACATATACAGAGTATATAGATATACCA 1937
Db 752 AATATATATATATATATATATATATATATATATATATATATATATATAT 694
QY 1938 CAACAACTTTTCTGTGATCAAAATATATATATATATATATATATAT 1995
Db 693 ATATATATATATATATATATATATATATATATATATATATATATATAT 634
QY 1996 TTCCATTTCCAGTACGTAACATTTACCTTAAATCGAATATCCATTTAT 2055
Db 633 TTATATATATATATATATATATATATATATATATATATATATATATAT 574
QY 2056 CGTGAGATGTATCAAAATTTCAATTTTATATATATATATATATATATAT 2115
Db 573 -TTAATATATATATATATATATATATATATATATATATATATATATAT 515
QY 2116 TCCTCTATGACCAAAATATAGAAAAACAAAAATATAGAACTATCCAA 2175
Db 514 AAATATATATATATATATATATATATATATATATATATATATATATAT 455
QY 2176 ACATTTTCACAATCATACAAAAACCTTT 2206
Db 454 ATATTTTATATATATATATATATATATATATATATATATATATATAT 424

```

```

RESULT 5
AG278124/c 1388 bp DNA linear GSS 18-DEC-2004
LOCUS Mus musculus molossinus DNA, clone:MSM91-048F07.77, genomic survey
DEFINITION sequence.
ACCESSION AG278124
VERSION AG278124.1 GI:47851001
KEYWORDS GSS.
SOURCE Mus musculus molossinus (Japanese wild mouse)
ORGANISM Mus musculus molossinus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

```

```

REFERENCE
AUTHORS
1 Abe, K., Noguchi, H., Tagawa, K., Yuzuriha, M., Toyoda, A., Kojima, T., Ezawa, K., Saitou, N., Hattori, M., Sakaki, Y., Moriwaiki, K. and Shiroishi, T.
TITLE
Contribution of Asian mouse subspecies Mus musculus molossinus to genomic constitution of strain C57BL/6J, as defined by BAC-end sequence-SNP analysis
JOURNAL
PUBMED 15574823
REFERENCE 2 (bases 1 to 1388)
AUTHORS Hattori, M., Toyoda, A., Noguchi, H., Kojima, T. and Sakaki, Y.
TITLE Direct Submission
JOURNAL Submitted (17-NOV-2003) Masahira Hattori, The Institute of Physical and Chemical Research (RIKEN), Genomic Sciences Center (GSC), Japan 1-7-22 Suenhiro-cho, Tsukuba, Ibaraki, 305-0074 Japan (E-mail: hattori@gsc.riken.jp, URL: http://hgp.gsc.riken.go.jp/, Tel: 81-45-503-9111, Fax: 81-45-503-9170)
COMMENT Clones are derived from the mouse BAC library MSM91. For BAC library availability, please contact Kuniya Abe (abe@rtc.riken.jp). The Institute of Physical and Chemical Research (RIKEN) 3-1-1 Koyadai, Tsukuba, 305-0074 Japan phone: 81-298-36-9189, fax: 81-298-36-9199 e-mail: abe@rtc.riken.jp
PRIMERS
Sequencing : T7
LIBRARY : pBACe3.6
VECTOR : EcoRI
R.site 1 : EcoRI
R.site 2 : EcoRI
FEATURES
source
location/Qualifiers
1..1388
/organism="Mus musculus molossinus"
/mol_type="genomic DNA"
/sub_species="molossinus"
/db_xref="taxon:57486"
/clone="MSM91-048F07.77"
/sex="male"
/tissue_type="mixture of kidney and spleen"
/clone_id="MSM91 Mouse Male BAC library"
ORIGIN
Query Match 3.9%; Score 87.2; DB 10; Length 1388;
Best Local Similarity 44.5%; Pred. No. 2e-06;
Matches 441; Conservative 0; Mismatches 539; Indels 11; Gaps 2;
QY 387 ATTATTTAAATGGAGAAATTTTGTGATACACCGTTAAGCAAGCATTAACAGCT 446
Db 1067 ATATTTATTAATTAATATATATATATATATATATATATATATATATATAT 1008
QY 447 AAAGTTACAACTTTAGAGATTTAGCAAAATTTACAAAGCTATATATACAGATTGA 506
Db 1007 ATTAATTAATTAATTAATTTGTAATTAATTAATTAATTAATTAATTAATTA 948
QY 507 TGATTTGAGAAATTTAAAAAGACTACAGCTCGGATTTACACATCATGAGCTTGA 566
Db 947 ATAGATATATATATTTAAATGATTAATTAATTAATTAATTAATTAATTA 889
QY 567 ACAAGTCGCTTGACTTTAAATATGATTTGAGAAATTTGACAAATGATTTATTCG 626
Db 888 TTAATATTAATTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 829
QY 627 AATACCTGCTTCAACTGAACTTATTAATTAATTAATTAATTAATTAATTAAT 686
Db 828 AAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 769
QY 687 TGCTAATTTTCAATTAATTTATTAACAAGAGTGTGAATTTGGCTGATGAAGGAT 746
Db 768 TAAACAAATTAATTTAAATTAATTTAAGAAATTAATTAATTAATTAATTAAT 709
QY 747 AGATATACATCTTCACAAATTTGAACCTAATGCTGAACATCAGATGATTAATTA 806

```

Db	708	ATTAATATATATTATTAATTATATATTTTATTAAGAAATTAATATATATATATATTT	649
QY	807	TTTAAAGAAATATATCTTAATATATAGTACATTTGTCAATACCTATAGACAGACT	866
Db	648	AGATTTAAAAAATATAAAAATTTTAAATAAATATAAAAATTAATATGTAATG	589
QY	867	AAAAATCTTAGACAGCAACCAATATGAATGAGTATTTATTAATGACTATCGAAGATA	926
Db	588	TAAATAAATATAAAAAAGTTAATATATTAATTAATTAATTAATTAATTAATTAATTA	529
QY	927	TATGACCTTACTGTATTTAGATACATCTTCATATTTCTTTATATGATATTAAGAATA	986
Db	528	AGAAAAAATTAATTAATAGAAAAATTAATTAATTAATTAATTAATTAATTAATTA	469
QY	987	TAGAGATTCATATAGAGAAATAGAAATGAAGTAAAGGCTTATAGATGAATCACAAGAGAAAT	1046
Db	468	TAAATTAAGTATATATTAAGGGGATATGATTTATTAATTAATTAATTAATTAATTAAT	409
QY	1047	TTATTAACAATGAATTAATTTTGTATGCTTCTTCACACTTATAGATTCAACCAATCTAGC	1106
Db	408	AAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT	349
QY	1107	TAGATGGAATATATATTTTAACACGTCGAAGTTTAAATATATTTCAATTTATAGAACATTT	1166
Db	348	AAGATTTTAAATATATATTAATTAATTAATTAATTAATTTGGTATATATTAATTAATTAATTA	289
QY	1167	TATTTTATTAACAGAAAAATACAAATTTTCGGAAATGTTTATGTTAGTATTTCTAATCGTA	1228
Db	288	TTTTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATG-	230
QY	1227	TGCACCTATATAGCAATATCTATATCTGAATCTTATATGAGAGAAAGACAGCTTACC	1286
Db	229	-----TAAATATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTT	179
QY	1287	CACAACAAAAACAATTAAGACCAATTTGAATCTTATTAAGTTTCAATTTGTACTGATAGACA	1346
Db	178	TAAATTTAAATATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA	119
QY	1347	ATCACTCTGTTTCCCTATTCACACACAC	1377
Db	118	CTCTCTCACAGGACACACACTCACACAC	88

RESULT	6
CNS0039G	
LOCUS	
DEFINITION	CNS0039G 1101 bp DNA linear GSS 03-JUN-1999
ACCESSION	Drosophila melanogaster genome survey sequence TERT end of BAC #
KEYWORDS	BACR08K10 of RPCT-98 library from Drosophila melanogaster (fruit fly), genomic survey sequence.
SOURCE	AL063921
ORGANISM	AL063921.1 GI:4941778
REFERENCE	GSS.
AUTHORS	Drosophila melanogaster (fruit fly)
JOURNAL	Drosophila melanogaster Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Epiphytoidae; Drosophilidae; Drosophila. 1 (bases 1 to 1101) Genoscope.
COMMENT	Direct Submission Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : sequef@genoscope.cns.fr Web : www.genoscope.cns.fr) Determination of this BAC-end sequence was carried out as part of a collaboration with the Berkeley Drosophila Genome Project (BDGP). The BDGP is constructing a physical map of the Drosophila melanogaster genome using these BACs. For further information please see <a href="http://www.fruitfly.org">http://www.fruitfly.org</a> The BDGP Drosophila melanogaster BAC library was prepared by Kazutoyo Osoegawa and Astron Mamoer in Pieter de Jong's laboratory in the Department of Cancer Genetics at the Roswell Park Cancer Institute in Buffalo, NY. The library is named RPCT-98 and was constructed by partial

Scorpi digestion of *Drosophila* DNA provided by the BDGP from the isogenic strain y2<sup>1</sup> cn bw sp, the same strain used for the BDGP P1 and BST libraries. A more detailed description of the library and how to order individual BAC clones, the entire library, or filters for hybridization from the BACPAC Resource center can be found at [http://bacpac.med.buffalo.edu/drosophila\\_bac.htm](http://bacpac.med.buffalo.edu/drosophila_bac.htm).

```
FEATURES
source      location/Qualifiers
1. 1101
/organism="Drosophila melanogaster"
/mol_type="genomic DNA"
/db_xref="taxon:7227"
/clone="BACR08X10"
/clone_1p="RPC1-98"
/note="end = TGT3"
```

## ORIGIN

Query Match	3.94	Score	66.8	DB	10	Length	1101
Best Local Similarity	19.14	Pred. No.	2.4e-06				
Matches	132	Conservative	289	Mismatches	261	Indels	8
						Gaps	1

[illegible]

RESULT 7	CG749499/c	LOCUS	CG749499	1348 bp	DNA	linear	GSS 24-OCT-2003
DEFINITION	P043-4-A06.za	Ppa	Scori	BAC Library	<i>Pristionchus pacificus</i>	genomic	c.



ACCESSION CG749499  
 VERSION CG749499.1 GI:37970425  
 KEYWORDS GSS  
 SOURCE Pristionchus pacificus  
 ORGANISM Pristionchus pacificus  
 Eukaryota; Metazoa; Nematoda; Chromadorea; Diplogasterida;  
 Neodiplogasteridae; Pristionchus.  
 1 (bases 1 to 1348)  
 Srinivasan,J., Sinz,W., Jesse,T., Wiggers-Perebolte,L., Jansen,K.,  
 Buntjer,J., van der Meulen,M. and Sommer,R.J.  
 An integrated physical and genetic map of the nematode Pristionchus  
 pacificus  
 JOURNAL Mol. Genet. Genomics 269 (5), 715-722 (2003)  
 PUBMED 12884007  
 COMMENT Contact: Sommer RJ  
 Evolutionary Biology  
 Max-Planck-Institute for Developmental Biology  
 Spemannstr. 37-39, Tuebingen D-72076, Germany  
 Tel.: 00497071601371  
 Fax: 00497071601498  
 Email: ralf.sommer@tuebingen.mpg.de  
 Class: BAC ends.

FEATURES  
 source  
 1..1348  
 Location/Qualifiers  
 /organism="Pristionchus pacificus"  
 /mol\_type="genomic DNA"  
 /strain="California"  
 /db\_xref="taxon:54126"  
 /clone\_lib="Ppa Ecoli BAC library"  
 /note="The library was generated by a partial digest of  
 the genomic DNA with Ecoli and cloning into the BAC  
 vector."

ORIGIN  
 Query Match 3.8%; Score 84.6; DB 10; Length 1348;  
 Best Local Similarity 43.1%; Pred. No. 6.4e-06;  
 Matches 506; Conservative 0; Mismatches 660; Indels 9; Gaps 2;

QY 390 TATTAAATGGAGAAATTTTGTGATACACCGTTAAACAGAAACATTAACAGCTTAA 449  
 DB TAT 1273  
 DB 1332 TAT 1273  
 QY 450 GTTACAACTTTAGAAAGATTAGCAATATTTACAAAGCTTAATACAGATTAGATGA 509  
 DB TAT 1213  
 DB 1272 AT 1213  
 QY 510 TTGGAGAAATTTAAAGACTACAGCTCTGATTTACACCATCATCAGATTACAACA 569  
 DB TTTAT 1153  
 DB 1212 TTTAT 1153  
 QY 570 AGTCGCCCTTGACTCTTAATATTCGATTGAGAAATTTGCAATGATTTTATTCGAAT 629  
 DB TTTTAAAT 1093  
 DB 1152 ATTTAAAT 1093  
 QY 630 ACCGCTTCCAACTTGAAGCTTAATAAGCTATTTACTACTATTTTATGCGCAAGCTGC 689  
 DB AT 1033  
 DB 1092 AT 1033  
 QY 690 TAAATTTCAATTAATTTATACACAAGTGTGATTTGCTGATGATGATGATGATGATG 749  
 DB AAAAT 973  
 DB 1032 AAAAT 973  
 QY 750 TATACATCTTCACAAATTTGAACCTTAATGCTGGAACATCAGATGCTATTTAACTTT 809  
 DB TAT 913  
 DB 972 AAAT 913  
 QY 810 AAAAGAAAT 869  
 DB TTTTAT 860  
 DB 912 AAAAT 860  
 QY 870 AAATCTTAGAGACGAACCAATATGAAATGAGTATATTTATGACTATCGAAGATATAT 929  
 DB TTTTAT 929

DB 859 AAAT 800  
 QY 930 GACCATTAAGTATATAGATACCATCTGCAATTTTCTTATATATATATATATATATATAG 989  
 DB TTTTAT 740  
 DB 799 AT 740  
 QY 990 AGATTCATATAGAGAAATATAGATATATATATATATATATATATATATATATATATAT 1049  
 DB AAAAT 680  
 DB 739 AAAAT 680  
 QY 1050 TACACAGTAAATATATTTGATCGTCTTCCATCAGCTTATGAGTCAACCAATCTAGCTAC 1109  
 DB AAAAT 620  
 DB 679 AAAAT 620  
 QY 1110 GATGGAAT 1169  
 DB AT 560  
 DB 619 AT 560  
 QY 1170 TTTTAT 1229  
 DB TTTTAT 502  
 DB 559 AAT 502  
 QY 1230 ACTTACTTATATGCAATCTATATATATATATATATATATATATATATATATATATAT 1289  
 DB TTTTAT 442  
 DB 501 AAAAT 442  
 QY 1290 AACAAAAACATATAGACATTTGATCTTATATATATATATATATATATATATATATAT 1349  
 DB TTTTAT 382  
 DB 441 TTTTAT 382  
 QY 1350 ACTCTCTGTTTCCCTTAT 1409  
 DB TTTTAT 322  
 DB 381 TTTTAT 322  
 QY 1410 TGCTCATCTTATACACACATCTATATATATATATATATATATATATATATATATAT 1469  
 DB TTTTAT 262  
 DB 321 ATTTAT 262  
 QY 1470 CACAACTTTTCTTAT 1529  
 DB TTTTAT 202  
 DB 261 AAAAT 202  
 QY 1530 TTCACAACTTTAT 1564  
 DB TTTTAT 167  
 DB 201 TTTTAT 167

RESULT 8  
 CDS00EVL 1101 bp DNA linear GSS 04-JUN-1999  
 LOCUS Drosophila melanogaster genome survey sequence T7 end of BAC:  
 DEFINITION BACR29B23 of Rpci-98 library from Drosophila melanogaster (fruit  
 fly), genomic survey sequence.  
 AL069706  
 ACCESSION AL069706.1 GI:4949849  
 VERSION  
 KEYWORDS GSS.  
 SOURCE Drosophila melanogaster (fruit fly)  
 ORGANISM Drosophila melanogaster  
 Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
 Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 Ephydroidea; Drosophilidae; Drosophila.  
 1 (bases 1 to 1101)  
 Genoscope.  
 REFERENCE Direct Submission  
 TITLE Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :  
 AUTHORS BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr  
 JOURNAL - Web : www.genoscope.cns.fr)  
 COMMENT Determination of this BAC-end sequence was carried out as part of a  
 collaboration with the Berkeley Drosophila Genome Project (BDGP).  
 The BDGP is constructing a physical map of the Drosophila  
 melanogaster genome using these BACs. For further information  
 please see <http://www.fruitfly.org> The BDGP Drosophila  
 melanogaster BAC library was prepared by Kazutoyo Osogawa and



[illegible]

	R Site 1	: Ecoref.
	R.Site 2	: Ecoref.
FEATURES	Location/Qualifiers	
source	. .1489	
	/organism="Mus musculus molossinus"	
	/mol_type="genomic DNA"	
	/sub_species="moiossinus"	
	/db_xref="taxon:57486"	
	/clone="MSMG01-146K14.TU"	
	/sex="male"	
	/tissue type="mixture of kidney and spleen"	
	/clone_lib="MSMG01 Mouse Male BAC Library"	
ORIGIN		
	Query Match	3.7%; Score 82.4; DB 10; Length 1489;
	Best Local Similarity	44.3%; Pred.No.1.7e=05;
	Matches	611; Conservative 0; Mismatches 734; Indels 33; Gaps 6
Dy	809 TAAAGAAAATATACTTAAATATAGTAACATTGTGCAATACCATTAGAACAAGACTPA	868
Dd	1487 TATATATATATTTAATAGTGTTTTATATATTTATATATTTATATTAATTAATATATTAATATGAATATAT	1428
Dy	869 AAAATCTTAGAGCAGAACCAAATATGGAATGGAGTAATTTTAATGACATCGGAACATTA	928
Dd	1427 AATATTAATATTTATTTAAAGNTATATTTAAATATTTATTAATATAATTAATANTGTGAAAAATTAATTA	1366
Oy	929 TGACCATTAACGTATTAAGTACCATTCCTCAATTTCTTATATATGATTTAAAAAGATATA	988
Dd	1367 TTATTATTATTAAGTATTAATATATATATATATATATATTAATTAATTAATATATATTAAT	1306

[illegible][illegible]

Oy		1642 ACACGACTAGCGTTGGATNGAATTAAATGAATACTCAAGAATAAATAATTAACAATGATCCCA	1701
Dd		664 -----TAAATTTAAATATATATATATATATAAAAATTAATTTTTATATATATATATATATTA	613
Oy		1702 GCAATCAAAGGTAAACAATCCTTGATTCACAAAACCTGAAGTAATGAAGAACCTGGTCATACA	1761
Dd		612 TTTAATTAATTAATATATATATATATATTTAAAAATTAATATATATTTAAATATATAAAAATATA	553
Oy		1762 GGAGGAAAATTGGITTAATTATTCAGAAAGTCAGGGCGTTTAGAAATTAACATGTGAAAACCTCT	1821
Dd		552 AATTTAATATATATATATATATATATATTTATTTATTTAAAAATTAATATTTATATATATATTT	493
Oy		1822 AATTCACACAACTTATTTTCCTTAGACTGCATATGCTACAAATAGCGCTCGAAATCATCT	1881
Dd		492 AAATATATATATATATATATATATATATATTAATAATGAATATTAATATATATATATATATTAAT	433
Oy		1882 CTTCCTTAATATATCTCTTACATATACCAGAGTAATATAGAAATCACCTCAACGACTCAAC	1941
Dd		432 TTAAATATATATATATTAABAAAAAAAAAATTTBABAATATATATATATATATATATTTAT	373
Oy		1942 AACACTTTTCTCGTACAAATATATATATTTCAATACGAGATTTGGGTATTTCCA	2001
Dd		372 ATTAATATTTTAT	313
Oy		2002 TTCCACAGTACAGTAACATTACCTTTAATGCAACATACACCTTATATTTAATGTCGA	2061
Dd		312 TAATA	253
Oy		2062 GATGTATCAAAATTCAAATTTTATATCATGTATATAAATTTGAAT--TTATACCAATATTCCTCTC	2120
Dd		252 AAATTAATA	193
Oy		2121 TAGGACCCAAATATAGAGAAAAACAATAATTAGAACTATGCCAAACAAAATTAATATCA	2178
Dd		192 TATATATATTTATTATAAAAATATTTTATTAATATATATTTAAAAAATTTATATATATATACA	135

RESULT 11  
CL110653/c

LOCUS CL110653 1594 bp DNA linear GSS 05-JAN-2004

DEFINITION ISB1-53P23\_Sp6.1 ISB1 Xenopus tropicalis genomic clone ISB1-53P23,  
genomic survey sequence.

ACCESSION CL110653

VERSION CL110653.1 GI:40604288

KEYWORDS GSS.

SOURCE Xenusus tropicalis (western clawed frog)

ORGANISM Xenusus tropicalis  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Amphibia; Batrachia; Anura; Mesobatrachia; Pipidea; Pipidae;  
Xenopodine; Xenusus; Silurana.  
1 (bases 1 to 1594)  
Kremitzki,C., Carter,J., McPherson,U., Warren,W., Graves,T.,  
Mardis,E. and Wilson,R.  
A physical map of the xenusus tropicalis genome  
Unpublished (2003)  
Contact: Richard K Wilson  
Genome Sequencing Center  
Washington University School of Medicine  
Email: rudmisions@wustl.wustl.edu  
Insert Length: 7500 Std Error: 0..00  
Seq primer: Sp6 ATTGAGTGACACATATAG  
Class: BAC ends  
High quality sequence start: 390  
High quality sequence stop: 470.  
Location/Qualifiers  

1..1594  
    <organism>"Xenusus tropicalis"  
    /mol\_type="genomic DNA"  
    /db\_xref="taxon:8364"  
    /clone="ISB1-53P23"  
    /clone\_1lb="ISB1"  
note=vector:pBelobacII; ISB-1 Xenusus tropicalis BAC

ORIGIN	Library Segment 1"
Query Match	3.7%; Score 82.4; DB 10; Length 1594;
Best Local Similarity	41.8%; Pred. No. 1.7e-05;
Matches	573; Conservative 0; Mismatches 790; Indels 8; Gaps 2
QY	362 AACACGACAAACAGTATGACACCAATTATTTAAATGGAGAAATTTTGTGATAC 421
DB	1414 AAAAAAAAAATTAATAAAAAAAAAAAAAAAAAAAAAAAAAATTAATTAATAATAA 1355
QY	422 CGTTAACGAGAGATTAACAGCTAAAGTTACAACTTTGAAAGATTGACCAATAT 481
DB	1354 TATATAATTAATTAATAAAAAAAAAAAAAAAAAAAAAAAAAATTAATAATAA 1295
QY	482 TACAAAGCTAATATACAGCATTTAGATGATGGAGAAATTAAGAACTACAGCTCTG 541
DB	1294 AAAAAAAAAAAAAAAAAATTAATAAAAAAAAAATTAATAATAATAATAATAATA 1233
QY	542 GATTACCAATCATCAGCATTTACACAGCTGCTTGACTTTAAATGCAATTGAGA 601
DB	1234 TAAAAAATAATTAATAAAAAAAAAAAAAAAAAATAATTAATAATAATAATAATA 1175
QY	602 ATGTTCAATGATTTTATTCGAGAAATCTGGTTCCACTTGAACCTTATAAAGC 661
DB	1174 AAAAAAAAAAACAATTAATAAAAAAAAAATTAATAATAATAATAATAATA 1115
QY	662 TATTACTACTATTTATGCGCAGCTGCTAATTTTCAATTTTATTTACACAGGTG 721
DB	1114 AAAAAAAAAAATTAATTAATAACCAAGAGAAAAAACAATAAAAAAAAAAAAA 1055
QY	722 CTGAATGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 781
DB	1054 ATAAATTAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 1001
QY	782 GAACATCGATGATTTATTAACCTTTTAAAGAAATATACCTAATATATGTAATT 841
DB	1000 AAATTAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 941
QY	842 GTGCATATACCTATAGAACAGACTAATAAATCTTAGACGAAACAATATGAAATGA 901
DB	940 TAAAAAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 881
QY	902 GTATATTTAATGACTATGAAATATATGACATTAATCTGATTTAGATACATCTCTCA 961
DB	880 AAAAAAAAAAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 821
QY	962 TTTCTTATATGATATTAAGATATGAGTTCAATGAGGAAATGAAATGAAGCA 1021
DB	820 AAAAAAAAAAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 761
QY	1022 TTAAGATGAACTCACAAGAGAAATTTATACACTGAATTAATTTGATGCTCTC 1081
DB	760 GAAAAAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 701
QY	1082 AACTTAGAGTTCAACCAATCTAGCTAGATGATATTAATTAAACGTCGAAGTTTA 1141
DB	700 ATATTAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 641
QY	1142 AATTAATTTCAATTTTACAAATTTATTTTATACGAAATATACAAATTTGGGAA 1201
DB	640 AAAAAAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 581
QY	1202 GTTTAGTTGATTTCTTAATCGTATGACCTACTTATAGCAATATCTTAATCTTA 1261
DB	580 AAAAAAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 521
QY	1262 TATATGAGAAAGAACGTTCAACCAACAATAAACAATATGATCTTATTA 1321
DB	520 AAAAAAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 461
QY	1322 AAGTTCAATTTGTAATGATGACATCACTCTGTTTCCCTAATTCACCACTTTA 1381



[illegible][illegible]







GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 11, 2005, 10:14:46 ; Search time 1318 Seconds  
(without alignments)  
1165.130 Million cell updates/sec

Title: US-10-783-417-1

Perfect score: 2208

Sequence: 1 atgatacaataacagataa.....atcatatacaaacacttga 2208

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 4996997 seqs, 3332346308 residues

Total number of hits satisfying chosen parameters: 9993994

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

N.Geneseq\_21:\*

- 1: geneseqn1980s:\*
- 2: geneseqn1990s:\*
- 3: geneseqn2000s:\*
- 4: geneseqn2001as:\*
- 5: geneseqn2001bs:\*
- 6: geneseqn2002as:\*
- 7: geneseqn2002bs:\*
- 8: geneseqn2003as:\*
- 9: geneseqn2003bs:\*
- 10: geneseqn2003cs:\*
- 11: geneseqn2003ds:\*
- 12: geneseqn2004as:\*
- 13: geneseqn2004bs:\*
- 14: geneseqn2005s:\*

Pred. NO. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2208	100.0	2208	13	ADR89394 AXMI-006
2	1903.8	86.2	2235	13	ADR89396 AXMI-007
3	1794.4	81.3	2085	13	ADR89398 AXMI-007
4	346.6	15.7	3940	1	AAV93054
5	346.6	15.7	4571	1	AAV93059
6	345	15.6	4934	1	AAV81490
7	340.6	15.4	3543	2	AAQ14669
8	340.6	15.4	3543	2	AAQ81178
9	188.2	8.5	2061	2	AAQ14670
10	188.2	8.5	2061	2	AAQ81180
11	178.2	8.1	3756	1	AAV50525
12	163.8	7.4	4896	6	ADQ43974
13	163.8	7.4	4896	10	ADP31301
14	163.8	7.4	4896	10	ADP31306
15	98.4	4.5	3690	6	ABK51132
16	94.6	4.3	15548	6	ABJ34155
17	90.4	4.1	8056	8	ABZ10246
18	89.6	4.1	2130	12	ADP71294
19	89.6	4.1	2246	12	ADP71295

20	86	3.9	4660	14	ADZ70902	Adz70902 Human mat
21	85.8	3.9	3444	2	AAV05249	AAV05249 CRYIF/cry
22	85.8	3.9	3444	2	AAV05251	AAV05251 CRYIF/cry
23	85.8	3.9	3444	2	AAV18723	AAV18723 CRYIF/cry
24	85.8	3.9	3444	2	AAV18701	AAV18701 CRYIF/cry
25	85.8	3.9	3444	2	AAV62080	AAV62080 Plasmid p
26	85.8	3.9	3444	2	AAV62082	AAV62082 Plasmid p
27	85.8	3.9	3450	2	AAV05269	AAV05269 CRYIA(c)/
28	85.8	3.9	3450	2	AAV18721	AAV18721 CRYIA(c)/
29	85.8	3.9	3450	2	AAV62079	AAV62079 Plasmid p
30	85.8	3.9	3522	2	AAQ10182	AAQ10182 Lepidopte
31	85.8	3.9	3522	2	AAQ47291	AAQ47291 Delta end
32	85.8	3.9	3522	2	AAV05250	AAV05250 CRYIF/436
33	85.8	3.9	3522	2	AAV05270	AAV05270 CRYIF/cox
34	85.8	3.9	3522	2	AAV18722	AAV18722 Codon-rev
35	85.8	3.9	3522	2	AAV18702	AAV18702 CRYIF/436
36	85.8	3.9	3522	2	AAV62083	AAV62083 Plasmid p
37	85.8	3.9	3522	2	AAV62081	AAV62081 Plasmid p
38	85.8	3.9	3522	2	AAV60606	AAV60606 Wild-type
39	85.8	3.9	3558	4	AAV00565	AAV00565 B. thurin
40	85.8	3.9	3558	4	AAV00420	AAV00420 B. thurin
41	85.8	3.9	4020	2	AAQ14648	AAQ14648 cryIF gen
42	82.2	3.7	1524	14	ADZ70897	ADZ70897 Human mat
43	81.2	3.7	6175	6	ABL33307	ABL33307 Human lmm
44	80	3.6	2482	14	ADZ71091	ADZ71091 Human chr
45	78.2	3.5	2131	14	ADZ71009	ADZ71009 Human chr

#### ALIGNMENTS

RESULT 1	ADR89394	standard; CDNA, 2208 BP.
ID	ADR89394	
XX	ADR89394;	
AC	18-NOV-2004 (first entry)	
XX		
DT		
XX		
DE	AXMI-006 coding sequence.	
XX		
KW	ss; gene; delta-endotoxin; delta-endotoxin associate polypeptide;	
KW	expression cassette; transformation; transgenic; plant; bacteria;	
KW	lepidoptera; coleoptera; pest; pesticide; resistance;	
KW	pesticidal activity.	
XX		
OS	Bacillus thuringiensis.	
XX		
FH	Key	Location/Qualifiers
FT	CDS	1..2208
FT		/*tag=a
FT		/product="AXMI-006"
PN	WO2004074462-A2.	
XX		
PD	02-SEP-2004.	
XX		
PF	20-FEB-2004; 2004WO-US005829.	
XX		
PR	20-FEB-2003; 2003US-0448632P.	
PR	20-FEB-2003; 2003US-0448633P.	
PR	20-FEB-2003; 2003US-0448797P.	
PR	20-FEB-2003; 2003US-0448806P.	
PR	20-FEB-2003; 2003US-0448810P.	
PR	20-FEB-2003; 2003US-0448812P.	
PR	19-FEB-2004; 2004US-00781979.	
PR	19-FEB-2004; 2004US-00782020.	
PR	19-FEB-2004; 2004US-00782036.	
PR	19-FEB-2004; 2004US-00782141.	
PR	19-FEB-2004; 2004US-00782570.	
PR	19-FEB-2004; 2004US-00783417.	
XX		
PA	(ATHE-) ATHENIX CORP.	



Db	1661	AAATAATTACAAATGATCCAGCAATCAAAGGTACAACTTTGATCAAACTCTAAGGTA	1740
Qy	1741	ATTGAAAGCACTGGTCATACAGAGAAACTTGGTTATTTTACAAAGTCAGAGGCGTTTA	1800
Db	1741	ATTGAAAGCACTGGTCATACAGAGAAACTTGGTTATTTTACAAAGTCAGAGGCGTTTA	1800
Qy	1801	GAAATTACATGTGAACCTCTTAATCTTACAACTTATTTTCATTAGACTTCATAGCT	1860
Db	1801	GAAATTACATGTGAACCTCTTAATCTTACAACTTATTTTCATTAGACTTCATAGCT	1860
Qy	1861	ACAATGTGTGTGAAATACCTCTTCATATATCTCTTACAATACCAAGAGTAATGGA	1920
Db	1861	ACAATGTGTGTGAAATACCTCTTCATATATCTCTTACAATACCAAGAGTAATGGA	1920
Qy	1921	ATACCACTCAACGACTCTCAACAAACCTTTTCTGGTCAAAATTATTAATTTTCAATAC	1980
Db	1921	ATACCACTCAACGACTCTCAACAAACCTTTTCTGGTCAAAATTATTAATTTTCAATAC	1980
Qy	1981	GGAGATTTTGGGATTTTCCAAATTTTCCAAAGTACAGTAACATTACCTTTAAATCGAAACATA	2040
Db	1981	GGAGATTTTGGGATTTTCCAAATTTTCCAAAGTACAGTAACATTACCTTTAAATCGAAACATA	2040
Qy	2041	CCATTATATTTAATTCGTGAGATGATGATCAAAATCAATTTAATCATTTGATTAATAATGGA	2100
Db	2041	CCATTATATTTAATTCGTGAGATGATGATCAAAATCAATTTAATCATTTGATTAATAATGGA	2100
Qy	2101	TTTATACCAATTACTTCTCTATGCAACCAAAATAGAGAAAAACAATAATTGAAACTATC	2160
Db	2101	TTTATACCAATTACTTCTCTATGCAACCAAAATAGAGAAAAACAATAATTGAAACTATC	2160
Qy	2161	CAAAACAAAATAATACATTTTTCACAAATCATACAAAACACCTTTGA 2208	
Db	2161	CAAAACAAAATAATACATTTTTCACAAATCATACAAAACACCTTTGA 2208	

19-FEB-2004; 2004US-00782096.  
PR 19-FEB-2004; 2004US-00782111.  
PR 19-FEB-2004; 2004US-00782570.  
PR 19-FEB-2004; 2004US-00783417.  
XX  
XX  
PA (ATHE-) ATHENIX CORP.  
XX  
XX Carozzi N, Hargiss T, Koziel MG, Duck NB, Carr B;  
XX WPI; 2004-635574/61.  
DR P-PSDB; ADR89397.  
XX  
PT New isolated delta-endotoxin and delta-endotoxin-associated nucleic acids  
PT and polypeptides, useful for killing lepidopteran or coleopteran pests or  
PT for producing organisms with pesticide resistance.  
PS  
PS  
PS Claim 1; SEQ ID NO 8; 178bp; English.  
XX  
XX This sequence encodes an isolated delta-endotoxin. Some of the delta-  
CC endotoxin coding sequences of the invention have alternative start  
CC codons, producing more than one protein from a single open reading frame.  
CC The nucleic acid sequences of the invention are useful in DNA constructs  
CC or expression cassettes for transformation and expression in plants and  
CC bacteria. The nucleic acids and corresponding polypeptides are useful for  
CC killing lepidopteran or coleopteran pests. Compositions containing the  
CC delta-endotoxins of the invention, and methods for their production, are  
CC useful for the production of organisms with pesticide resistance,  
CC specifically bacteria and plants. These organisms are useful for  
CC generating altered or improved delta-endotoxin or delta-endotoxin-  
CC associated proteins that have pesticidal activity, or for detecting the  
CC presence of delta-endotoxin or delta-endotoxin-associated proteins or  
CC nucleic acids in products or organisms.  
XX  
XX Sequence 2235 BP; 861 A; 364 C; 316 G; 694 T; 0 U; 0 Other;

RESULT 2  
ADR89396  
ID ADR89396 standard; cDNA; 2235 BP.

AC ADR89396;

DT 18-NOV-2004 (first entry)

DE AXMI-007 coding sequence

ss; gene; delta-endotoxin; delta-endotoxin associate polypeptide; KW

KW lepidoptera: pest; pesticide: resistance:

pesticidal activity.

XX  
05 Bacillus curingiensis.  
XX

EH	Key	Location/Qualifiers
ET	CDS	1 2235

```
FT      / * Lag = a
ET      / no adjustment
```

```
FT      /transl_except= pos:1. .3, aa:Met
vv
```

PN WO2004074462-A2.

PD 02-SEP-2004.

PF 20-FEB-2004; 2004WO-US005829.

PR 20-FEB-2003; 2003US-0448632P.

PR 20-FEB-2003; 2003US-0448797P.

PR 20-FEB-2003; 2003US-0448810P.

PR 19-FEB-2004; 2004US-00781979.

1  
 2  
 3  
 4  
 5  
 6  
 7  
 8  
 9  
 10  
 11  
 12  
 13  
 14  
 15  
 16  
 17  
 18  
 19  
 20  
 21  
 22  
 23  
 24  
 25  
 26  
 27  
 28  
 29  
 30  
 31  
 32  
 33  
 34  
 35  
 36  
 37  
 38  
 39  
 40  
 41  
 42  
 43  
 44  
 45  
 46  
 47  
 48  
 49  
 50  
 51  
 52  
 53  
 54  
 55  
 56  
 57  
 58  
 59  
 60  
 61  
 62  
 63  
 64  
 65  
 66  
 67  
 68  
 69  
 70  
 71  
 72  
 73  
 74  
 75  
 76  
 77  
 78  
 79  
 80  
 81  
 82  
 83  
 84  
 85  
 86  
 87  
 88  
 89  
 90  
 91  
 92  
 93  
 94  
 95  
 96  
 97  
 98  
 99  
 100  
 101  
 102  
 103  
 104  
 105  
 106  
 107  
 108  
 109  
 110  
 111  
 112  
 113  
 114  
 115  
 116  
 117  
 118  
 119  
 120  
 121  
 122  
 123  
 124  
 125  
 126  
 127  
 128  
 129  
 130  
 131  
 132  
 133  
 134  
 135  
 136  
 137  
 138  
 139  
 140  
 141  
 142  
 143  
 144  
 145  
 146  
 147  
 148  
 149  
 150  
 151  
 152  
 153  
 154  
 155  
 156  
 157  
 158  
 159  
 160  
 161  
 162  
 163  
 164  
 165  
 166  
 167  
 168  
 169  
 170  
 171  
 172  
 173  
 174  
 175  
 176  
 177  
 178  
 179  
 180  
 181  
 182  
 183  
 184  
 185  
 186  
 187  
 188  
 189  
 190  
 191  
 192  
 193  
 194  
 195  
 196  
 197  
 198  
 199  
 200  
 201  
 202  
 203  
 204  
 205  
 206  
 207  
 208  
 209  
 210  
 211  
 212  
 213  
 214  
 215  
 216  
 217  
 218  
 219  
 220  
 221  
 222  
 223  
 224  
 225  
 226  
 227  
 228  
 229  
 230  
 231  
 232  
 233  
 234  
 235  
 236  
 237  
 238  
 239  
 240  
 241  
 242  
 243  
 244  
 245  
 246  
 247  
 248  
 249  
 250  
 251  
 252  
 253  
 254  
 255  
 256  
 257  
 258  
 259  
 260  
 261  
 262  
 263  
 264  
 265  
 266  
 267  
 268  
 269  
 270  
 271  
 272  
 273  
 274  
 275  
 276  
 277  
 278  
 279  
 280  
 281  
 282  
 283  
 284  
 285  
 286  
 287  
 288  
 289  
 290  
 291  
 292  
 293  
 294  
 295  
 296  
 297  
 298  
 299  
 300  
 301  
 302  
 303  
 304  
 305  
 306  
 307  
 308  
 309  
 310  
 311  
 312  
 313  
 314  
 315  
 316  
 317  
 318  
 319  
 320  
 321  
 322  
 323  
 324  
 325  
 326  
 327  
 328  
 329  
 330  
 331  
 332  
 333  
 334  
 335  
 336  
 337  
 338  
 339  
 340  
 341  
 342  
 343  
 344  
 345  
 346  
 347  
 348  
 349  
 350  
 351  
 352  
 353  
 354  
 355  
 356  
 357  
 358  
 359  
 360  
 361  
 362  
 363  
 364  
 365  
 366  
 367  
 368  
 369  
 370  
 371  
 372  
 373  
 374  
 375  
 376  
 377  
 378  
 379  
 380  
 381  
 382  
 383  
 384  
 385  
 386  
 387  
 388  
 389  
 390  
 391  
 392  
 393  
 394  
 395  
 396  
 397  
 398  
 399  
 400  
 401  
 402  
 403  
 404  
 405  
 406  
 407  
 408  
 409  
 410  
 411  
 412  
 413  
 414  
 415  
 416  
 417  
 418  
 419  
 420  
 421  
 422  
 423  
 424  
 425  
 426  
 427  
 428  
 429  
 430  
 431  
 432  
 433  
 434  
 435  
 436  
 437  
 438  
 439  
 440  
 441  
 442  
 443  
 444  
 445  
 446  
 447  
 448  
 449  
 450  
 451  
 452  
 453  
 454  
 455  
 456  
 457  
 458  
 459  
 460  
 461  
 462  
 463  
 464  
 465  
 466  
 467  
 468  
 469  
 470  
 471  
 472  
 473  
 474  
 475  
 476  
 477  
 478  
 479  
 480  
 481  
 482  
 483  
 484  
 485  
 486  
 487  
 488  
 489  
 490  
 491  
 492  
 493  
 494  
 495  
 496  
 497  
 498  
 499  
 500  
 501  
 502  
 503  
 504  
 505  
 506  
 507  
 508  
 509  
 510  
 511  
 512  
 513  
 514  
 515  
 516  
 517  
 518  
 519  
 520  
 521  
 522  
 523  
 524  
 525

Query Match	86.2%	Score 1903.8;	DB 13;	Length 2235;
Best Local Similarity	92.2%;	Pred. No. 0;		
Matches 2030; Conservative	0;	Mismatches 162;	Indels 9;	Gaps 2

[illegible]

2  
 3  
 4  
 5  
 6  
 7  
 8  
 9  
 10  
 11  
 12  
 13  
 14  
 15  
 16  
 17  
 18  
 19  
 20  
 21  
 22  
 23  
 24  
 25  
 26  
 27  
 28  
 29  
 30  
 31  
 32  
 33  
 34  
 35  
 36  
 37  
 38  
 39  
 40  
 41  
 42  
 43  
 44  
 45  
 46  
 47  
 48  
 49  
 50  
 51  
 52  
 53  
 54  
 55  
 56  
 57  
 58  
 59  
 60  
 61  
 62  
 63  
 64  
 65  
 66  
 67  
 68  
 69  
 70  
 71  
 72  
 73  
 74  
 75  
 76  
 77  
 78  
 79  
 80  
 81  
 82  
 83  
 84  
 85  
 86  
 87  
 88  
 89  
 90  
 91  
 92  
 93  
 94  
 95  
 96  
 97  
 98  
 99  
 100  
 101  
 102  
 103  
 104  
 105  
 106  
 107  
 108  
 109  
 110  
 111  
 112  
 113  
 114  
 115  
 116  
 117  
 118  
 119  
 120  
 121  
 122  
 123  
 124  
 125  
 126  
 127  
 128  
 129  
 130  
 131  
 132  
 133  
 134  
 135  
 136  
 137  
 138  
 139  
 140  
 141  
 142  
 143  
 144  
 145  
 146  
 147  
 148  
 149  
 150  
 151  
 152  
 153  
 154  
 155  
 156  
 157  
 158  
 159  
 160  
 161  
 162  
 163  
 164  
 165  
 166  
 167  
 168  
 169  
 170  
 171  
 172  
 173  
 174  
 175  
 176  
 177  
 178  
 179  
 180  
 181  
 182  
 183  
 184  
 185  
 186  
 187  
 188  
 189  
 190  
 191  
 192  
 193  
 194  
 195  
 196  
 197  
 198  
 199  
 200  
 201  
 202  
 203  
 204  
 205  
 206  
 207  
 208  
 209  
 210  
 211  
 212  
 213  
 214  
 215  
 216  
 217  
 218  
 219  
 220  
 221  
 222  
 223  
 224  
 225  
 226  
 227  
 228  
 229  
 230  
 231  
 232  
 233  
 234  
 235  
 236  
 237  
 238  
 239  
 240  
 241  
 242  
 243  
 244  
 245  
 246  
 247  
 248  
 249  
 250  
 251  
 252  
 253  
 254  
 255  
 256  
 257  
 258  
 259  
 260  
 261  
 262  
 263  
 264  
 265  
 266  
 267  
 268  
 269  
 270  
 271  
 272  
 273  
 274  
 275  
 276  
 277  
 278  
 279  
 280  
 281  
 282  
 283  
 284  
 285  
 286  
 287  
 288  
 289  
 290  
 291  
 292  
 293  
 294  
 295  
 296  
 297  
 298  
 299  
 300  
 301  
 302  
 303  
 304  
 305  
 306  
 307  
 308  
 309  
 310  
 311  
 312  
 313  
 314  
 315  
 316  
 317  
 318  
 319  
 320  
 321  
 322  
 323  
 324  
 325  
 326  
 327  
 328  
 329  
 330  
 331  
 332  
 333  
 334  
 335  
 336  
 337  
 338  
 339  
 340  
 341  
 342  
 343  
 344  
 345  
 346  
 347  
 348  
 349  
 350  
 351  
 352  
 353  
 354  
 355  
 356  
 357  
 358  
 359  
 360  
 361  
 362  
 363  
 364  
 365  
 366  
 367  
 368  
 369  
 370  
 371  
 372  
 373  
 374  
 375  
 376  
 377  
 378  
 379  
 380  
 381  
 382  
 383  
 384  
 385  
 386  
 387  
 388  
 389  
 390  
 391  
 392  
 393  
 394  
 395  
 396  
 397  
 398  
 399  
 400  
 401  
 402  
 403  
 404  
 405  
 406  
 407  
 408  
 409  
 410  
 411  
 412  
 413  
 414  
 415  
 416  
 417  
 418  
 419  
 420  
 421  
 422  
 423  
 424  
 425  
 426  
 427  
 428  
 429  
 430  
 431  
 432  
 433  
 434  
 435  
 436  
 437  
 438  
 439  
 440  
 441  
 442  
 443  
 444  
 445  
 446  
 447  
 448  
 449  
 450  
 451  
 452  
 453  
 454  
 455  
 456  
 457  
 458  
 459  
 460  
 461  
 462  
 463  
 464  
 465  
 466  
 467  
 468  
 469  
 470  
 471  
 472  
 473  
 474  
 475  
 476  
 477  
 478  
 479  
 480  
 481  
 482  
 483  
 484  
 485  
 486  
 487  
 488  
 489  
 490  
 491  
 492  
 493  
 494  
 495  
 496  
 497  
 498  
 499  
 500  
 501  
 502  
 503  
 504  
 505  
 506  
 507  
 508  
 509  
 510  
 511  
 512  
 513  
 514  
 515  
 516  
 517  
 518  
 519  
 520  
 521  
 522  
 523  
 524  
 525  
 526

Year	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022	2023	2024	2025	2026	2027	2028	2029	2030	2031	2032	2033	2034	2035	2036	2037	2038	2039	2040	2041	2042	2043	2044	2045	2046	2047	2048	2049	2050	2051	2052	2053	2054	2055	2056	2057	2058	2059	2060	2061	2062	2063	2064	2065	2066	2067	2068	2069	2070	2071	2072	2073	2074	2075	2076	2077	2078	2079	2080	2081	2082	2083	2084	2085	2086	2087	2088	2089	2090	2091	2092	2093	2094	2095	2096	2097	2098	2099	2100
1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022	2023	2024	2025	2026	2027	2028	2029	2030	2031	2032	2033	2034	2035	2036	2037	2038	2039	2040	2041	2042	2043	2044	2045	2046	2047	2048	2049	2050	2051	2052	2053	2054	2055	2056	2057	2058	2059	2060	2061	2062	2063	2064	2065	2066	2067	2068	2069	2070	2071	2072	2073	2074	2075	2076	2077	2078	2079	2080	2081	2082	2083	2084	2085	2086	2087	2088	2089	2090	2091	2092	2093	2094	2095	2096	2097	2098	2099	2100	

[illegible]

131 CAGGATACCACTAAGTGGTCAGATTGTGT CAGCGGAAACACCAAATATTTGATAAT TCC 13

DB 128 CAAATTAACAAAGAGTGGCTCAATATATGTCACAAAGGAAATACACAAATATGGTGAATAAATTCG 18

191 AGACATTTGCTAGTCTGATACAAATTGCTGCAGTTAGTGCAGTACTATGTATCCGTA 25

Db 188 AGACATTTCCTAGTCTGATACATTGCTGCAGTTAGTGCAGGTACTATTGATCCGTA 24

251 CTCTGTTAGCCGGTATAGGTGGCTCACTTCTATATCCGACCGATAGGAATAATAGTC 31

Db 248 CTCTGTTAGCCGGTATAGGTGGGCTCACTTCTATATCCGGACCGATAGGAATAATAGGTG 30

311 CTATATAATCTTTGGTACCCATACTGCTTTTGGCCGGGAGACAAGACA 37

Db 308 CTATAATAATCTTTGGTACCCCTAATCACTGTCTTTGGCCCGGGGAGACAAGACA 36

371 AAACAGTATGGACACAACTTATTAAATGGGAAATTTTGTGATACACCGTTAACAG 43

Db 368 AAACAGTATGCACAACTTATTAAATGGGAAATTTTGTGATACACCGTTAACAG 42

431 AAAGCATAAACAGCTAAAGTTACAACCTTTAGAAGGATTAGACAATAATTACAAGCT 49

Db 428 AAAGCATAAACAGCTAAAGTTACAACTTTAGAGGATTTAGACAAATATTACAAGCT 48

491 ATAATACAGCATTAGATGATTGGAGAAATTAAGACTACAGCTCCTGGATTACCAC 55

\_\_\_\_\_

Db 488 ATAAATACAGCATAGATGATGGAGAAAAATTAAAAAGCTACAGCTCTGGATTACAC 547  
 Qy 551 CATCATCAGCATTTACACAGAGCTGCTTGACTCTTAAAAATACGATTTGAGAAATGTTACA 610  
 Db 548 CATCATCAGCATTTACACAGAGCTGCTTGACTCTTAAAAATACGATTTGAGAAATGTTACA 607  
 Qy 611 ATGATTTATTCGAGAAATACCTGGTTTCCAACTTGAACCTTAATAACCGTATTACTAC 670  
 Db 608 ATGATTTATTCGAGAAATACCTGGTTTCCAACTTGAACCTTAATAACCGTATTACTAC 667  
 Qy 671 CTATTTATGCCAGCTGCTTAATTTCAATTAAATTTTACACAGAGTCTGAATGG 730  
 Db 668 CTATTTATGCCAGCTGCTTAATTTCAATTAAATTTTACACAGAGTCTGAATGG 727  
 Qy 731 CTGATGAATGGAATGCAATATACCTCTTCCAAATTTGAACCTTAATGCTGGAACATCAG 790  
 Db 728 CTGATGAATGGAATGCAATATACCTCTTCCAAATTTGAACCTTAATGCTGGAACATCAG 787  
 Qy 791 ATGACTATTAATACTTTTAAAAAGAAATATACCTAAATATATGTAACCTATTGTGCAATA 850  
 Db 788 ATGACTATTAATACTTTTAAAAAGAAATATACCTAAATATATGTAACCTATTGTGCAATA 847  
 Qy 851 CCTATAGAACAGAGCTTAAAAATCTTAGAGAGAACCAATATGAAATGAGATATTATTA 910  
 Db 848 CCTATAGAACAGAGCTTAAAAATCTTAGAGAGAACCAATATGAGATGAGATATTATTA 907  
 Qy 911 ATGACTATGAGATATATGACCAATTAATGATATGATACCAATCTCTCAATTTTCTTTAT 970  
 Db 908 ATGATTTATTCGAGAAATATGACTATTACTATGATATGATACCAATCTCTCAATTTTCTTTAT 967  
 Qy 971 ATGATTTAAAAAGATATATGATATGATATGATATGATATGATATGATATGATATGATATG 1030  
 Db 968 ATGATTTAAAAAGATATATGATATGATATGATATGATATGATATGATATGATATGATATG 1021  
 Qy 1031 AACTCACAAGAGAAATTTTATCAACTGAAATTAATTTTATGATGCTCTCAACTTATAG 1090  
 Db 1022 AACTCACAAGAGAAATTTTATCAACTGAAATTAATTTTATGATGCTCTCAACTTATAG 1081  
 Qy 1091 TTCAACCCCAATCTGATGATGATGATATTAATTAACAGTGCAGAGTTTAAATTTAT 1150  
 Db 1082 TTCAACCCCAATCTGATGATGATGATATTAATTAACAGTGCAGAGTTTATTTAT 1141  
 Qy 1151 CATTTTGAACAATTTTATTTTATATAGAAAAATTCGAGAAATGCTTATAGTTG 1210  
 Db 1142 CATTTTGAACAATTTTATTTTATATACAAAAAAGAAACGAGGGAATGCTTATAGTTG 1201  
 Qy 1211 GTATTTCTAATCGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1270  
 Db 1202 GTATTTCTAATCGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1261  
 Qy 1271 AAAGAACAGGTTCAACCAACAAACAAATAGACCAATTTGAATCTTATTAAGTTTCAA 1330  
 Db 1262 AAAGAACAGGTTCAACCAACAAACAAATTTGAATCTTATTAAGTTTCAA 1321  
 Qy 1331 TTGTATCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1387  
 Db 1322 TTGTATCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1381  
 Qy 1388 ATCAAAATGAACTTATTTTAAATGCTGATCAACACACACTCAATATTCAGAGAG 1447  
 Db 1382 ATCAAAATGAACTTATTTTAAATGCTGATCAACACACACTCAATATTCAGAGAG 1441  
 Qy 1448 GGTCTTTATCTAATATGCAAAACAGAACTTTTTCAAATTTCTGAAAAAAGACTGCA 1507  
 Db 1442 GGAATTTATCTAATATGCAAAACAGAACTTTTTCAAATTTCTGAAAAAAGACTGCA 1501  
 Qy 1508 ATCTAGTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1567  
 Db 1502 AACCAATTTATTAATCAAAATGTTTACCAAGCTAATAATGATATTTATATCC 1561  
 Qy 1568 ATTTTCTATTTATTTCTATTTCTATGATGATGATGATGATGATGATGATGATGATGATGATG 1627  
 Db 1562 AGTTTCTTTATTTATTTATTTCTATTTATTTATTTATTTATTTATTTATTTATTTATTTAT 1621

Qy 1628 GTGTATTTAGATGACACACAGTATGTTGATGATATTAATGCAATATCAGATTAATA 1687  
 Db 1622 GTGTATTTAGATGACACACAGTATGTTGATGATATTAATGCAATATCAGATTAATA 1681  
 Qy 1688 TTCAATGATCCCGAGCAATCAAGGTAAACAATCTGTATCAAACTCTAAGGTAATTGA 1747  
 Db 1682 TTCAATGATCCCGAGCAATCAAGGTAAACAATCTGTATCAAACTCTAAGGTAATTGA 1741  
 Qy 1748 GACCTGTATACAGAGAGAACTTGGTTTATTTATTAACAAGTCAAGGCGTTTGAATA 1807  
 Db 1742 GACCTGTATACAGAGAGAACTTGGTTTATTTATTAACAAGTCAAGGCGTTTGAATA 1801  
 Qy 1808 CATGTAAATCTCTAATTTCTACACATCTTATTTCAATTAGACTTGATGATCAAAATG 1867  
 Db 1802 CATGTAAATCTCTAATTTCTACACATCTTATTTCAATTAGACTTGATGATCAAAATG 1861  
 Qy 1868 GTGCTGGAATATCTCTCTAATATATCTCTTCAATACAGAGATATGGAATATACAC 1927  
 Db 1862 GTGCTGGAATATCTCTCTAATATATCTCTTCAATACAGAGATATGGAATATACAC 1921  
 Qy 1928 CTCAACGACTCAACACACTTTTCTGTACAAATTTATTAATTTTCAATACGAGATT 1987  
 Db 1922 CTCAACGACTCAACACACTTTTCTGTACAAATTTATTAATTTTCAATACGAGATT 1981  
 Qy 1988 TTGGGTAATTTCCAAATTTCCAGTACAGTAACTTATCTTTAATCGAAACATACATT 2047  
 Db 1982 TTGGGTAATTTCCAAATTTCCAGTACAGTAACTTATCTTTAATCGAAACATACATT 2041  
 Qy 2048 TATTTATCGTACAGATGATCAAAATCAATTTTATCATTTGATTAATGAAATTTATAC 2107  
 Db 2042 TATTTATCGTACAGATGATCAAAATCAATTTTATCATTTGATTAATGAAATTTATAC 2101  
 Qy 2108 CAATTTACTCTCTATGACCAACAAATAGAGAAAAAATTAAGAACTATCCAAACAA 2167  
 Db 2102 CAATTTACTCTCTATGACCAACAAATAGAGAAAAAATTAAGAACTATCCAAACAA 2161  
 Qy 2168 AAATTAATCAATTTTTCACAAATCATCAAAACACCTTTGA 2208  
 Db 2162 AAATTAATCAATTTTTCACAAATCATCAAAACACCTTTGA 2202

RESULT 3  
 ADR89398  
 ID ADR89398 standard; cDNA; 2085 BP.  
 XX  
 AC ADR89398;  
 XX  
 DT 18-NOV-2004 (first entry)  
 XX  
 DE AXMI-007 alternative start site coding sequence.  
 XX  
 KW ss; gene; delta-endotoxin; delta-endotoxin associate polypeptide;  
 KW expression cassette; transformation; transgenic; plant; bacteria;  
 KW lepidoptera; coleoptera; pest; pesticide; resistance;  
 KW pesticial activity.  
 XX  
 OS Bacillus thuringiensis.  
 XX  
 FH Key location/Qualifiers  
 FT 1..2085  
 FT CDS /product= "Alternative AXMI-007"  
 XX  
 PD MO2004074462-A2.  
 XX  
 PD 02-SEP-2004.  
 XX  
 PF 20-FEB-2004; 2004MO-US005829.  
 XX  
 PR 20-FEB-2003; 2003US-0448632P.  
 PR 20-FEB-2003; 2003US-0448633P.  
 PR 20-FEB-2003; 2003US-0448797P.

20-FEB-2003; 2003US-0448806P.  
PR 20-FEB-2003; 2003US-0448810P.  
PR 20-FEB-2003; 2003US-0448812P.  
PR 19-FEB-2004; 2004US-00781979.  
PR 19-FEB-2004; 2004US-00782020.  
PR 19-FEB-2004; 2004US-00782096.  
PR 19-FEB-2004; 2004US-00782141.  
PR 19-FEB-2004; 2004US-00782570.  
PR 19-FEB-2004; 2004US-00783417.  
PA (ATHE-) ATHENIX CORP.  
XX Carozzi N, Hargis T, Koziel MG, Duck NB, Carr B;  
XX WPI; 2004-635574/61.  
DR P-PSDB; ADR89399.  
XX  
XX New isolated delta-endotoxin and delta-endotoxin-associated nucleic acids  
PT and polypeptides, useful for killing lepidopteran or coleopteran pests or  
PT for producing organisms with pesticide resistance.  
XX  
XX Claim 1; SEQ ID NO 10; 178bp; English.  
XX  
XX This sequence encodes an isolated delta-endotoxin. Some of the delta-  
CC endotoxin coding sequences of the invention have alternative start  
CC codons, producing more than one protein from a single open reading frame.  
CC The nucleic acid sequences of the invention are useful in DNA constructs  
CC or expression cassettes for transformation and expression in plants and  
CC bacteria. The nucleic acids and corresponding polypeptides are useful for  
CC killing lepidopteran or coleopteran pests. Compositions containing the  
CC delta-endotoxins of the invention, and methods for their production, are  
CC useful for the production of organisms with pesticide resistance.  
CC Specifically bacteria and plants. These organisms are useful for  
CC generating altered or improved delta-endotoxin or delta-endotoxin-  
CC associated proteins that have pesticidal activity, or for detecting the  
CC presence of delta-endotoxin or delta-endotoxin-associated proteins or  
CC nucleic acids in products or organisms.  
XX  
XX Sequence 2085 BP; 794 A; 338 C; 302 G; 651 T; 0 U; 0 Other;  
SQ  
Query Match 81.3%; Score 1794.4; DB 13; Length 2085;  
Best Local Similarity 92.7%; Pred. No. 1.9e-302;  
Matches 1908; Conservative 0; Mismatches 141; Indels 9; Gaps 2;  
QY 154 ATGTCTCAAGGGAATACCAATATGTGTATTTGAGACATTTGCTAGTGTATCA 213  
DB 1 ATGTCTCAAGGGAATACCAATATGTGTATTTGAGACATTTGCTAGTGTATCA 60  
QY 214 ATGTCTGAGTAGTGACAGTACTATTTGATCCGGTACTCTGTAGCCGTATAGTGG 273  
DB 61 ATGTCTGAGTAGTGACAGTACTATTTGATCCGGTACTCTGTAGCCGTATAGTGG 120  
QY 274 CTGACTTGTATATCCGGACCGATAGGAATATATGTGTATATATATCTTTTGGTACC 333  
DB 121 CTGACTTGTATATCCGGACCGATAGGAATATATGTGTATATATATCTTTTGGTACC 180  
QY 334 CTAAATCATCTGTTTGGCCCGCGGAGAAACAAGCAAGTATGACCAATTTAT 393  
DB 181 CTAAATCATCTGTTTGGCCCGCGGAGAAACAAGCAAGTATGACCAATTTAT 240  
QY 394 AAAATGGAGAAATTTTGTGTATACCGTTAACAAGAGCATTAACAGTTAAAGTTA 453  
DB 241 AAAATGGAGAAATTTTGTGTATACCGTTAACAAGAGCATTAACAGTTAAAGTTA 300  
QY 454 CAACTTTAGAGATTTAGACAAATATTTAACAAGCTATATATACAGATTAGATTTGG 513  
DB 301 CAACTTTAGAGATTTAGACAAATATTTAACAAGCTATATATACAGATTAGATTTGG 360  
QY 514 AGAAATTTAAAGATCAAGCTCTGTATTAACAACCATCATCAGCATTAAGAAGT 573  
DB 361 AGAAATTTAAAGATCAAGCTCTGTATTAACAACCATCATCAGCATTAAGAAGT 420  
QY 574 GCCTTGACTTTAAATAGATTTGAGATTTTCAATGATTTTATTCGAGAAATACCT 633

|||||  
DB 421 GCCTTGACTTTAAATAGATTTGAGATTTTCAATGATTTTATTCGAGAAATACCT 480  
QY 634 GGTTCACACTTGAACCTTATTAACGCTATTAACCTATTTTATGCGCAAGCTGAT 693  
DB 481 GGTTCACACTTGAACCTTATTAACGCTATTAACCTATTTTATGCGCAAGCTGAT 540  
QY 694 TTTCAATTTAAATTTATTAACAAGAGTCTGAATTTGCTGATGATGATGATGATGAT 753  
DB 541 TTTCAATTTAAATTTATTAACAAGAGTCTGAATTTGCTGATGATGATGATGATGAT 600  
QY 754 CATCTTCACAAAATTTGAACCTTAATGCTGGAACATCATGATGATGATGATGATGAT 813  
DB 601 CATCTTCACAAAATTTGAACCTTAATGCTGGAACATCATGATGATGATGATGATGAT 660  
QY 814 GAAATATATACCTTAATATATGATGATGATGATGATGATGATGATGATGATGATGAT 873  
DB 661 GAAATATATACCTTAATATATGATGATGATGATGATGATGATGATGATGATGATGAT 720  
QY 874 CTTAGAGACGAACCAATATGATGATGATGATGATGATGATGATGATGATGATGATGAT 933  
DB 721 CTTAGAGACGAACCTTAATATGATGATGATGATGATGATGATGATGATGATGATGAT 780  
QY 934 ATTAAGTATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 993  
DB 781 ATTAAGTATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 840  
QY 994 TCAATAGAGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1053  
DB 841 TCAATAGAGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 894  
QY 1054 ACTGAAATATATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1113  
DB 895 ACTGAAATATATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 954  
QY 1114 GAATATATATTTTACACGTCGATGATGATGATGATGATGATGATGATGATGATGATGAT 1173  
DB 955 GAATATATATTTTACACGTCGATGATGATGATGATGATGATGATGATGATGATGATGAT 1014  
QY 1174 TATACAGAAATATCAAAATTTGGGGAATGCTTATGCTTATTTTATTCGATGATGATGAT 1233  
DB 1015 TATACAGAAATATCAAAATTTGGGGAATGCTTATGCTTATTTTATTCGATGATGATGAT 1074  
QY 1234 ACTTATAGCAATATCAATATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1293  
DB 1075 ACTTATAGCAATATCAATATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1134  
QY 1294 AAAACATATAGCAATATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1353  
DB 1135 AAAACATATATCAATATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1194  
QY 1354 CCGTGTTCCTGATTC--AACACACTTATATATTAATTAATTAATTAATTAATTAAT 1410  
DB 1195 CCGTGTTCCTGATTC--AACACACTTATATATTAATTAATTAATTAATTAATTAAT 1254  
QY 1411 GGCCTCATATTAACAACATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1470  
DB 1255 AATTCACCTAGTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1314  
QY 1471 ACAACCTTTTTCATTTTCTAGAAAAAAGACTGCAATCTAGTATTTATGATGATGATGAT 1530  
DB 1315 ACAACCTTTTTCATTTTCTAGAAAAAAGACTGTAACAATTAATTAATTAATTAATTAAT 1374  
QY 1531 TCACCAACCTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1590  
DB 1375 TCACCAACCTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1434  
QY 1591 TATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1650  
DB 1435 TATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1494  
QY 1651 AGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1710

```
D 1495 AGTGTATAGAAATATGCAATATCAGATAAAATTAATTAACATGATCCAGCAATCAA 1554
Q 1711 GGTAACAATCTTGATGACAACTTAAGTAATGGAAGACCTGGTCATACAGAGGAAAC 1770
D 1555 GGTAACGCTCTGATACCAACTCTTAAGTAATGGAAGACCTGGTCATACAGAGGAAAC 1614
Q 1771 TTGGTTATTTACAAAGCAAGGCGTTTGAAGAAATTAATGGAAGACCTCTTAATTTTCA 1830
D 1615 TTGGTTATTTACAAAGCAAGGCGTTTGAAGTAATGATTAATGATGAACTCTTAATTTTCA 1674
Q 1831 CAATCTTATTTATGATGATCTGATATGCTCAAAATGGTGTGGAATATCTCTTCAAT 1890
D 1675 CAATCTTATTTATGATGATCTGATATGCTCAAAATGGTGTGGAATATCTCTTCAAT 1734
Q 1891 ATATCTTTACAAATTCAGAGTAATAGAAATACCACTCAACGACTCAACAACTTTT 1950
D 1735 ATATCTTTACAAATTCAGAGTAATAGAAATACCACTCAACGACTCAACAACTTTT 1794
Q 1951 TCTGTTACAAATTTATTAATTTTACAAATGCGAGATTTTGGGTATTTCCAAATTTCCAA 2010
D 1795 TCTGTTACAAATTTATTAATTTTACAAATGCGAGATTTTGGGTATTTCCAAATTTCCAA 1854
Q 2011 ACAGTAACATTTACCTTTAAATCGAAACATACCATTTATTTAAATCGTACAGATGATCA 2070
D 1855 ACAGTAACATTTACCTTTAAATCGAAACATACCATTTATTTAAATCGTACAGATGATCA 1914
Q 2071 AATTCATTTTATTCATTTGATTAATTTGAATTTATTCATTTACTCTCTGATGACCA 2130
D 1915 AATTCATTTTATTCATTTGATTAATTTGAATTTATTCATTTACTCTCTGATGACCA 1974
Q 2131 AATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2190
D 1975 AATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2034
Q 2191 CATACAAAGACCTTTGA 2208
D 2035 CATACAAAGACCTTTGA 2052

RESULT 4
AAN93054
ID AAN93054 standard; DNA; 3940 BP.
XX
AC AAN93054;
XX
DT 27-AUG-2003 (revised)
DT 25-MAR-2003 (revised)
DT 30-MAY-1990 (first entry)
XX
DB Delta-endotoxin crystal protein gene.
XX
KW Delta-endotoxin; crystal protein; insecticide; ss; pcc130;
KW biological control agent.
XX
OS Bacillus thuringiensis; israeliensis.
XX
FH Key Location/Qualifiers
FT RBS 879..884
FT /tag= b
FT /label= Shine-Delgarno sequence
FT CDS 891
FT /*tag= a
FT /product= "delta-endotoxin"
XX
EP296870-A.
XX
PD 28-DEC-1988.
XX
PF 24-JUN-1988; 88EP-00305772.
XX
PR 26-JUN-1987; 87US-00067653.
XX
PA (DUPO ) DU PONT DE NEMOURS & CO E I.
```

```
XX
FI Ellar DU, Ward ES;
XX
DR WPI; 1989-001322/01.
XX
PT P-PSDB; AAP94035.
XX
PT DNA fragment encoding insecticidal protein - obtd. from Bacillus
FT thuringiensis sub species israeliensis, and used in microorganisms and
PT plant cells.
XX
PS Disclosure; Fig 5; 26pp; English.
XX
CC The nucleotide sequence is an insert in plasmid pcc130. The delta
CC endotoxin protein is insecticidal and can be used to control insect pests
CC esp. mosquitoes. See also AAN93059. (Updated on 25-MAR-2003 to correct PA
CC field.) (Updated on 27-AUG-2003 to correct OS field.)
XX
SQ Sequence 3940 BP; 1457 A; 603 C; 659 G; 1221 T; 0 U; 0 Other;

Query Match 15.7%; Score 346.6; DB 1; Length 3940;
Best Local Similarity 53.2%; Pred. No. 6e-51;
Matches 1176; Conservative 0; Mismatches 934; Indels 99; Gaps 17;

Q 4 AATCAAAATTAACGTAATTAAGAAATTAATGATTCGCACTACCTCACTATTTT 63
D 888 AATATGAAATCTTATCAAAATTAATGAAATGAAATCAATTAATGCTTCAAAAAAAA 947
Q 64 CCGAACAAGAAACAGTAATGATTTAGATACCTTACACAAATATCCAAATCAACATTA 123
D 948 TTAATATATCTAATATTAATTAACAGATATCCAAATATGTCACAAATTAATTA 1007
Q 124 CAAACACAAATTAACAAAGAGTGGCTCAATATGTCACAGGAATTAACAATATGTCAT 183
D 1008 CAAAGTCAAAATTAATTAAGATTTGGCTCAATATGTCACAGGAATTAACAATATGTCAT 1067
Q 184 AATTGAGACATTTGCTAGTGCATGATCAATTCGTCAGTGTGAGGACTATTTGA 243
D 1068 GATTTGAAACTTTATATGAT-----AGTGTAACTCAGTGTCTTACTATTTGA 1118
Q 244 TCCGTAATCTGTTAGCCGGTATAGTGGGCTCACTTATATCCGACCGATAGATA 303
D 1119 GTTGGACCGTACTGACTGGT-----TTCCGGTTCACAACACC 1157
Q 304 ATAGTGTCTATTAATATATCTTTTGGTACCTTAATCACTGTCTTTGGCCCGGAGAA 363
D 1158 TTAGGACTTCTTAAATAGGTTTGGTATTAATATCAATGATCTTTTCCAGCCCAAGAC 1217
Q 364 CAAGACAAACAGATAGGACACAAATTAATTAATAGGAGAAATTTTGTGATACACCG 423
D 1218 CAATCTAACAC--ATGGAAGTACTTTATTAACACAAATTAATTAATTAATTAATTA 1274
Q 424 TTAAACA-GAAGCATTAACAGCTTAAGTTACAACTTTAGAGATTTAGACAAATATT 482
D 1275 ATAGCATCAACATATTAATAGTAATGCTAATTAATTTTAAACAGTGGTTTAAATGTTATC 1334
Q 483 ACAAGCTATTAATTAACAGCATTAATGATGATGAGAGAAATTAATAAGACTCAACGCTCTCG 542
D 1335 AGCACTTATATATATCACTTAAACATGAGGAGATATATCCAAACCAAAATATCTCAG 1394
Q 543 ATTACGACCATCATCAGCATTAACAAGAGTGGCTTACCTTAAATATAGATTTGAGAA 602
D 1395 GATGTAAGGA-----CACAAATCCAGCTAGTTCATTAACATTTTCAAAATGTCATTCACA 1450
Q 603 TGTTCACAATGATTTTATTCGAGAAATATCCGTGTTTCAACTGGAACCTTAATAACGCT 662
D 1451 GCTTGAATACCTTGTCTCTTAATCTTAATGATGTC-----GATTAATTAACATTA 1504
Q 663 ATTACTACTATTTATTAAGCAAGTGTCTAAATTTTCAATTAATTAATTAACAACAAGTGC 722
D 1505 AGTATTAATCTAGTATACCAAGCAAGCAAACTTAATCTAGTATTAATTAATCAAGCGGT 1564
Q 723 TGAATTGGCTGATGAATGGAATGCAATATCAATCTTCAAAATTAAGAACTTAATGCTGG 782
```

```

Db 1565 CAAATTTGAAGGTTATTTAAAAACAATCGACAAATTCGATTTATTTAGACCTTTGCC--- 1621
Oy 783 AACATCAGATGCTATTTATTAACCTTTTAAAGAAAATATACCTAAATATATGATCTATG 842
Db 1622 AACGCAATTTGATTTATTTCCAGTATTTGCTAAAGCTATAGAAATTTACATATATATG 1681
Oy 843 TCGAAATATCTATAGAACAGACTAAAAATCTTAGACGACAAACCAATATGAAATGAG 902
Db 1682 TGTAAACAATTATATAAAAAGATTAAATTTATTAACGACCGCTGATAGTAATCTTGA 1741
Oy 903 TATATTTATGACTATCGAAGATATATGACCTATCTGATTTATGATACCATTCTCAATT 962
Db 1742 TGGAAATTTAACTCGAACACATA-----CAATCGATTCGACAAATAATCACTACTGC 1795
Oy 963 TTCTTTATATGATATATAAAGATATAGATTCATATAGAGGATATAGAAAGGAT 1022
Db 1796 TGTATTTAATGTTGTTGCACTCTTCTTATATATGATGAGGTAATATCCAAATGAGT 1855
Oy 1023 TAAGATGAACTCACAAGAAATTTATATCACTGAAATTAATTTGATCTCTCA 1082
Db 1856 CCAATCTGAATCTACGAGAAATTTAT--CAGTACTTAATCTTGAGAAAGCCCTAT 1913
Oy 1083 ACTTGAAGTTCAACCAATCTAGCTACGATGGAATATATTTAACAGTCAAGTTTAA 1142
Db 1914 AAATATATATGACTTTCATATCAGAGATTCATTAACGTAAGACC--GCATTTATTTA 1971
Oy 1143 ATATTTTCAATTTTATAGAACATTTATTTTATACAGAAATATCAAAATTTGGGAATCG 1202
Db 1972 CTGGCTGATCTTTGAA-----TTTATATGAAAAAGGCAAACTTCTCTAATATAT 2026
Oy 1203 TTTAGTTGATTTCTTAAATCGATGACCTTATATAGCAATATATATGAAACTTT 1262
Db 2027 TTTTACACGCAATTTATATATGTTTATTTACACACTGTATATATATCCAAAAATCTAG 2086
Oy 1263 ATATGAGAAAGAACAGGTTCAACCAACAAAATATAGACATTTGAACTTTATTA 1322
Db 2087 TGTATTTGAAATCACAATATGATATATTAATAATCTCTGTTGGCAAAATAT 2146
Oy 1323 AGTTTCAATTTGATCTGATAGACATCACTCCCTTTCCCTATTTCAACCACTTTAT 1382
Db 2147 TTTATATTTTATTTTAAATGTCATAGCTTATATATATATATATGATTTAATTA 2206
Oy 1383 AATTATCAAAATTTGAATTTTATTAATGCTCTATTAACAACACACTCAATATTTAGC 1442
Db 2207 TATTTAGTAAATGATTTTATTTATTAATGATCTAGACTTTTGGAAAAAGACTTAC 2266
Oy 1443 AGAGAGGCTCT-----TATCTAATTTCAAAAACAACCTTTTTCATTTCTAGAAA 1496
Db 2267 AGCAGGATCTGGGCAAAATTAATCTTATGATGTAAATATAATTTTTCGGGTTTCAATCT 2326
Oy 1497 AAAAGACTGCAATCTAGTATTTGATCCAGGTTTTCACCAAACTTTAATTAATATGTC 1556
Db 2327 TAAACGAAAGAGAAATCAAGAAACCTTACCTTTTCCAAATATGATTAATATGTC 2386
Oy 1557 TATTTTATCCATTTTTCATTTACTTATTTCTATGATGATTTGATTTACAGCTCAAT 1616
Db 2387 TATTTTATC-----ATTATTAATAAGCTTTAGTATCCCTGCAACATAT 2429
Oy 1617 ATTAGATACAGGTGA--TTAGATGACACACACTAGTGTGATATATATATGCAATA 1674
Db 2430 AAAACTCAAGTATACGTTGCTTGACACACTAGTGTGATCTTAATAATATCAAT 2489
Oy 1675 TCAGATTAATAATATACATATGCCCAATCAAGATTAACATCTTGATCAAACT 1734
Db 2490 TATACACTTTTACTACCAAAATTCAGCTGTAAGGAATTCACCTGGGCTGCTCT 2549
Oy 1735 AAGGTAATGAGAGCTGTCATACAGAGGAACTTGTTATTTTCAAAAGTCAGAG 1794
Db 2550 AAGGTTGTTCAAGAGCTGTCATACAGAGGGAATTTATTTGATTTCAAAAGATCA---- 2605
Oy 1795 CGTTTGAATTAATCAATGTAATCTCTTAATTTCTAACAACTTTATTTAGACTTGA 1854
Db 2606 -TTTCAAAATTAACATGTCACACTCAAAATTTTCAACAACTGATTTATTAAGAAATCGT 2663

```

```

Oy 1855 TATGCTACAAATGCTGTGAAATATCTCTCTAATATATCTTACATACAGAGTA 1914
Db 2664 TATGCTTCAATATGAAAGGCAATATCTGAGCTGTTATTAATCTTATATCCAGGG-- 2721
Oy 1915 ATAGAAATACCACTCAACGACTCAACAAACCTTTTCTGTAGCAATTAATTAATTTA 1974
Db 2722 -TAGCAAGATCGGGATATGCACTCAACCCCACTTTTCTGGTACGATTTATGCAATTTA 2780
Oy 1975 CAATACGAGATTTTGGGTATTTCCAAATTTCCAAATGATACAGTAATTTACTTTAATGCA 2034
Db 2781 AAATATTAAGATTTTCACTACTAGAAATTTCTAACGAGGTAATTTGCTCAAAATCAA 2840
Oy 2035 AACATACATTTATATTTATATCTGTCAGATGAT--CAATTTCAATTTTATCAATGAT 2091
Db 2841 AACATATCTCTTGTTGTTTATATGTTTGAATGATATATACAAACACAGTACTTATGAT 2900
Oy 2092 AAAATGAAATTTATACCAATTAATCTCTCTATGACCAAAATATAGAAAAACAAATTTA 2151
Db 2901 AAAATGAAATTTCTGCAATTAATCTGTTCTATTAAGAGAGATATAGAGAAACAAATTTA 2960
Oy 2152 GAACTATCCAAACAAATTAATTAATCAATTTTTCACAAATTCATACAAAA 2200
Db 2961 GAAACAGTACAAACAAATTAATTAATCAATTTATGCAAAATCCTATTAATAA 3009

RESULT 5
AAN93059
ID AAN93059 standard; DNA; 4571 BP.
XX
AC AAN93059;
XX
DT 27-AUG-2003 (revised)
DT 25-MAR-2003 (revised)
DT 30-MAY-1990 (first entry)
XX
DE Delta-endotoxin crystal protein gene.
XX
KM Delta-endotoxin; crystal protein; insecticide; pCH130; see:
  biological control agent.
XX
OS Bacillus thuringiensis; israelensis.
XX
FH Key Location/Qualifiers
FT RBS 879..884
FT 1/tag= b
FT 2/tag= d
FT 3/tag= Shine-Dalgarno sequence.
FT misc_feature 891..4430
FT 1/tag= a
FT 2/product= "delta-endotoxin crystal protein"
XX
EN EP296870-A.
XX
PD 28-DEC-1988.
XX
PF 24-JUN-1988; 88EP-00305772.
XX
PR 26-JUN-1987; 87US-00067653.
XX
PA (DUPO ) DU PONT DE NEMOURS & CO B I.
XX
PI Eljar DJ, Ward ES;
XX
DR WPI, 1989-001322/01.
DR P-PsDB; AAP93715.
XX
PT DNA fragment encoding insecticidal protein - obcd, from Bacillus
  thuringiensis sub species israelensis, and used in microorganisms and
  plant cells.
XX
PS Disclosure; Fig 7; 26pp; English.
XX
CC The sequence encodes the 130 kDa delta-endotoxin gene from B.

```

cc thuringiensis subsp. israeliensis and is a 4.46 kb insert of pCH30. The  
cc delta endotoxin protein is insecticidal and can be used to control insect  
cc pests esp. mosquitoes. See also AN93054. (Updated on 25-MAR-2003 to  
cc correct PA field.) (Updated on 27-AUG-2003 to correct OS field.)  
xx

SO Sequence 4571 BP; 1678 A; 685 C; 817 G; 1391 T; 0 U; 0 Other;

Query Match 15.7%; Score 346.6; DB 1; Length 4571;  
Best Local Similarity 53.2%; Pred. No. 6e-51;  
Matches 1176; Conservative 0; Mismatches 934; Indels 99; Gaps 17;

```
OY 4 AATCAAAATTAACGAATATACGAATATGAAATTTATGATTCGATPACTGCTATTTT 63
DB AATATGAAATCTTATCAAAATTAATTAATGAAATGAAACATTTAAATGCTTCAAAAAA 947
OY 64 CCGAAGCAAGAAACAGTATGATTTAGATACCTTTACACAAATTAATCCAAATCAACATTA 123
DB TTAATATATCTAATTAATTAATCAAGATATCAATAGAAATATGCCAAAACATTTATTA 1007
OY 124 CAAAACAAATTTACAAAGAGTGGCTCAATATGTCTCAAGGGAATACAAATATGTGTAT 183
DB CAAAGTCAAAATTAATTAAGATTTGGCTCAATATGTCTCAAGAAATCAAGACATATGTGA 1067
OY 184 AATTTCAGACATTTGCTAGTGTGATACAAATTCGTGAGTTAGTCAAGTACTATTGTA 243
DB GATTTTGAACCTTTATATGAT-----AGTGTGAACCTGAGTCTTATATCTATTTGA 1118
OY 244 TCCGTAATCTGTTAGCCGTAATAGTGGGCTCACTTATATCCGACCGATAGATATA 303
DB GTTGGGACCGTACTGACTGT-----TTCCGGTTTCAACACACC 1157
OY 304 ATTAGTCTATTAATTAATATCTTTTGGTACCTTAATCACTGTCTTTTGGCCCGGGAGAA 363
DB TTAGGACTGTCTTAATATAGTTTGGTATCAATTAATACAGTTCTTTTCCAGCCCAAGAC 1217
OY 364 CAAGCAAAACAGATATGACAAATTTATTAATTAATGGAATTTTGTGATATACCG 423
DB CAATCTAACAC--ATGAGTGAATCTTAATTAACAAACTAAATTAATTAATTAATTAATTA 1274
OY 424 TTAACA--GAAGATTAATTAATGATTAAGTTTCAAACTTTAGAGATTTAGACAAATTT 482
DB ATAGATCAACATTAATTAATGATTAATGATTAATTAATTAATTAATTAATTAATTAATTA 1334
OY 483 ACAAGCTATTAATTAATGATTAATGATTAATGATTAATTAATTAATTAATTAATTAATTA 542
DB AGCATTAATCAATTAATCACTTAATTAATGAGATTAATTAATTAATTAATTAATTAATTA 1394
OY 543 ATTACCAATCATCATGATTAATCAACAGTGCCTTGAATCTTAATTAATTAATTAATTA 602
DB GATGTAAAGG---CAAAATCCAGCTAGTCAATTAATTAATTAATTAATTAATTAATTAATTA 1450
OY 603 TGTTCACATGATTTTATTTGCAAAATTAATCTGTGTTCCAACTTGAATTTAAACGCT 662
DB GCTTGTAATCTGTGCTCTCTCAATCTTAATTAATTAATTAATTAATTAATTAATTAATTA 1504
OY 663 ATTACTACTTATTTATTTGCGCAAGCTGCTAATTTTCAATTTAAATTTATTAATTAATTA 722
DB AGTATTAATCTATTAATTAATGCAAGGCAAGCAACTTCACTGATTAATTAATTAATTAATTA 1564
OY 723 TGAATTTGCTATGAATGGAATGCAATATATCAATCTTCAAAATTAATTAATTAATTA 782
DB CAATTTTAAGCGTATTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1621
OY 783 AACATCAATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 842
DB AATGCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1681
OY 843 TGCATTAATCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 902
DB TGTAACTAACTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1741
OY 903 TATATTTAATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 962
```

```
DB 1742 TCGAATATTAATCTGGAACACATA-----CAATACGTATCGAACAAAATGACTACTGC 1795
OY 963 TTCTTTATGATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1022
DB TGTATTAATGATTTGTGCACTCTTCTCTTAATTAATTAATTAATTAATTAATTAATTAATTA 1855
OY 1023 TAGAATGAATCTCAAGAGAAATTTTATACACTGAATTAATTAATTTGATGCTCTTCTCA 1082
DB CCAATCTGAACCTTAATCTGGAATTTAT--CAGGTACTTAATCTTGAAGAAAGCCCTAT 1913
OY 1083 ACTTAGAGTTCAACCCATCTAGCTAGATGAATATTAATTAATTAATTAATTAATTAATTAATTA 1142
DB AATATTAATGATCTTCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1971
OY 1143 ATTAATTTCAATTTTATGAACAAATTTATTTTATTAACAGAAATTAATCAAAATTTGGGAAATCG 1202
DB CTGCGCTTGAATCTTTGAA-----TTTTTATGAAGAAAGCCGAACCTCTTAATTAATTT 2026
OY 1203 TTTAGTTGATTTTCTAATGCTGATGCACTTACTTAATAGCAATCTTAATCTGAACCTTT 1262
DB TTTTACCAAGCATTAATTAATTAATTTTCAATCACTGATTAATTAATTAATTAATTAATTA 2086
OY 1263 ATATGAGAAAGAACAGGTTCAACCAACAAACAAATTAAGACATTTGATCTTTAA 1332
DB TGTTTTGAATATCAATGTAATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2146
OY 2087 AGTTTCAATTTGATCTGATAGCAATCACTCTGTTTCCCTAATTAATCAACACATTTAT 1382
DB TTATATTTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2206
OY 1383 AATTATCAATTTGAATTTATTTAATTTGCTCATCTTAACAACTCAATTAATTTACG 1442
DB TATTAGTAATTAATGATTTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2266
OY 1443 AGGAGGCT-----TTATCTAATTAATCAAAACAACTTTTTCGAATTTCTTAGAAA 1496
DB AGCAGATCTGGGCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2267
OY 1497 AAAAGCTGCAATCTAGTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1556
DB TAAACGAAGAGAAATCAAGAAACCTTACCCCTTTTCAACATTAATTAATTAATTAATTAATTA 2386
OY 1557 TATTTATCCCATTTTCAATTTATTTACTTAATTTCTATATGATTTGATTTACAGCTTCAAT 1616
DB TATTTATTC-----ATTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2429
OY 1617 ATTAGATACAGGTGA--TTAGATGGAACACAGTATGTTGATTAATTAATTAATTAATTA 1674
DB AAAATCAAGTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2489
OY 1675 TCAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1734
DB TATACATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2549
OY 1735 AAGTATTTGAAGACCTGCTCATTAACAGAGAACTTTGTTTATTTTCAAAAGTCAAGG 1794
DB AAGGTTTTCGAAGACCTGCTCATTAACAGAGAGGATTTTATTTGATTTTCAAAAGATCA---- 2605
OY 1795 CGTTTGAATTTCAATGTAAGAACTCTTAATTTTCAACAACTTTTCAATTTTCAATTTTCA 1854
DB --TTTCAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2606
OY 1855 TATGCTCAAAATGCTGTAAGAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1914
DB TATGCTTCAAAATGTAAGAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2721
OY 1915 ATAGGAATTAACCTCAAGCTCAACAACTTTTCTGCTGATTAATTAATTAATTAATTAATTA 1974
DB -TAGCAAGACCTGCTGATTAAGCACTCAACCTTTTCTGCTGATTAATTAATTAATTAATTAATTA 2780
OY 1975 CAATAGGAGATTTTGGTATTTTCCAAATTTTCAAGTACAGTAACTTAATTAATTAATTAATTA 2034
DB AATATTAAGATTTTCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2781
DB 2781 AATATTAAGATTTTCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2840
```



QY 2035 AACATACATTATATTTATGTCGAGATGAT---CAATTCATTTTATCATTTGAT 2091  
 DB 2841 AACATATCTCTTGTGTTTATGTTGATGATATATACAAACAGTACTTATTTGAT 2900  
 QY 2092 AAAATGTAATTTATACCAATTAATCTCTTATGACCAAAATAGAGAAAACAAAATTA 2151  
 DB 2901 AAAATGTAATTTCTGCAATTAATCTCTTATAGAGAGATAGAGAGAAAAATTA 2960  
 QY 2152 GAAACTATCCAAACAAATAATACATTTTCACAAATCATACAAAA 2200  
 DB 2961 GAAACGATACAAATAATTAATACATTTTATGCAAAATCTTATAAAA 3009  
 RESULT 6  
 AAN81490  
 ID AAN81490 standard; DNA; 4934 BP.  
 XX AAN81490;  
 AC  
 XX 24-OCT-2003 (revised)  
 DT 05-NOV-1990 (first entry)  
 DE Insecticidal (Diptera) protein gene.  
 XX Insecticidal protein gene; Diptera; Aedes; ss.  
 KM Insecticidal protein gene; Diptera; Aedes; ss.  
 OS Bacillus thuringiensis serovar israelensis.  
 XX  
 FH Key Location/Qualifiers  
 FT CDS 1..4908  
 FT /\*tag= a  
 FT /product= "insecticidal protein"  
 FT 620..1368  
 FT /\*tag= b  
 FT intron  
 PN JP63230090-A.  
 XX 26-SEP-1988.  
 PD  
 XX 19-MAR-1987; 87JP-00066844.  
 PF  
 XX 19-MAR-1987; 87JP-00066844.  
 PR  
 XX 19-MAR-1987; 87JP-00066844.  
 PA (SUMO) SUMITOMO CHEM IND KK.  
 XX  
 DR WPI; 1988-311968/44.  
 DR P-PSDB; AAP81034, AAP82589.  
 XX  
 PT New insecticidal protein of bacillus thuringiensis israelensis strain -  
 PT prepred. by isolating plasmid contg. insecticidal protein gene by forming  
 PT gene library from plasmid deoxyribonucleic acid, etc.  
 XX  
 PS Disclosure; Page 7; 9pp; Japanese.  
 CC A plasmid contg. the gene is isolated by forming a gene library from  
 CC plasmid DNA of B. thuringiensis israelensis (HD 522 strain, USA Goldberg  
 CC ONR0) followed by screening with anti-Israelensis insecticidal protein  
 CC IG. The insecticidal protein is highly effective against Diptera, esp.  
 CC Aedes. (Updated on 24-OCT-2003 to standardise OS field)  
 CC  
 XX Sequence 4934 BP; 1797 A; 749 C; 877 G; 1511 T; 0 U; 0 Other;  
 SQ  
 Query Match 15.6%; Score 345; DB 1; Length 4934;  
 Best Local Similarity 53.2%; Pred. No. 1.1e-50;  
 Matches 1175; Conservative 0; Mismatches 935; Indels 99; Gaps 17;  
 QY 4 AATCAAAATTAACGATTAATGAAATATGAAATTAATGATGATACATACCTTATTTT 63  
 DB 1366 AATATGAATCTTATCAAAATTAATGAAATGAAATTAATGCTTACAAAAAAA 1425  
 QY 64 CCGAACAGAAACGATTAATGATTTCTAGATACCTTACACAAATTAATCAATCAACATTA 123

DB 1426 TTAATATATCTAATTAATTAACAAGATATCCAAATAGAAAAATAGTCCAAACATTAATTA 1485  
 QY 124 CAAAACAAATTAACAAAGAGTGGCTCAATATAGTCAAGGAATATACAAATATGATGAT 183  
 DB 1486 CAAAGTAAATTAATTAAGATTTGGCTCAATATAGTCAAGGAATATAGATATGATGAT 1545  
 QY 184 AATTCGAGACATTTGCTAGTCTGATCAATTCGTCGAGTTAGTGCAGATTAATTTGA 243  
 DB 1546 GATTTGAAATCTTTATGAT-----AGTGTGAATCTAGTCCATATCTATTTGTA 1596  
 QY 244 TCCGCTACTCTGTAGCCGGATATAGTGGGCTCACTTATATCCGAGCCATATAGAA 303  
 DB 1597 GTTGGACCGTATGACTGT-----TTGGGTTTCAACACCCC 1635  
 QY 304 ATAGTGTCTAATTAATTAATCTTTTGTACCTTATCATCTGTCTTTTGGCCGCGGAGAA 363  
 DB 1636 TTAGACTTGTCTTATATACGTTTGTGATCATTAATACAGATTTCTTTTCCAGCCCAAGAC 1695  
 QY 364 CAAGCAAAACAGATATGACACAAATTTTAAATGGAGAAATTTTGTGATACACCG 423  
 DB 1696 CAATCTAACAC---ATGAGATGACTTTATTAACAAACCTAAATATTAATTAATTAAGAA 1752  
 QY 424 TTAAAC-GAAAGCATTAACAGCTAAAGTACAACTTTAGAGGATTTAGCAAAATTT 482  
 DB 1753 ATAGCATCAATATATATAGTATGCTAATTAATTTAAACAGGTCTTTAATGTTATC 1812  
 QY 483 ACAAAGCTATATATCAGATTAATGATTTAGAGAAATTTAAAGACTACAAAGCTCCTGG 542  
 DB 1813 AGCACTTATCATTAATACCTTAAACATGAGAGAAATATCCAAACCCCAAAATATCTGAG 1872  
 QY 543 ATTACCAACCATCATCAGCATTAACAAACAGCTGCTTGAATCTTAAATATACGATTTGAGA 602  
 DB 1873 GATGTAAAGA---CAGAAATCCAGCTATGTTATTCATTAATCAATTTTCAAAATGTCATTCAGA 1928  
 QY 603 TGTTCACATATATTTATTTAGAGAAATACCTGTTTCCAACTTGAACCTTAATAAACGT 662  
 DB 1929 GCTTGTAAACTCTGTCTCTTAATCTTATGATGATTC-----GATTAATATATATAT 1982  
 QY 663 ATTACTACTTATTAATGCGAAGCTGCTAATTTTCAATTTAATTAATTAACAAGAGTGC 722  
 DB 1983 AGTATTAATCTATGATTAATGACAAAGCAAGCAACTTATCATCTGATGATTAATTAATCAAGCGT 2042  
 QY 723 TGAATTCGCTGATGATGAATGACATATATCATCTTACAAATTTGAACCTAATGCTGG 782  
 DB 2043 CAATTTGAACGCTATTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2099  
 QY 783 AACATCAGATATCTATTAATTAATTTAAAGAAATATACCTTAATTAATTAATTAATTA 842  
 DB 2100 AACAGCAATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2159  
 QY 843 TGCAAAATACCTTAATGAAGAGACTTAATTAATTTAGAGACGAAACCAATTAATGAATGAG 902  
 DB 2160 TGTAAACAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2219  
 QY 903 TATATTTATGATCATGATGATATATGATGATGATGATGATGATGATGATGATGATGAT 962  
 DB 2220 TGGAAATTAATCTGGAACACATA-----CAATACGATGAGCAAAATGATCACTACG 2273  
 QY 963 TTTCTTATATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1022  
 DB 2274 TGTATTAATCTTGTGTGACCTTTCTTAATTAATTAATTAATTAATTAATTAATTAATTA 2333  
 QY 1023 TTAAGAAATGATCAAGAGAAATTTTAATTAATTAATTAATTAATTAATTAATTAATTA 1082  
 DB 2334 CCAATCTGAACCTTACTGAGAAATTTAT--CAGGTACTTAATCTTGAAGAAAGCCCTAT 2391  
 QY 1083 ACTTGAAGTTCAACCCATCTAGCTAGATGAATATTAATTAATTAATTAATTAATTAATTA 1142  
 DB 2392 AATATTAATGATCTTCAATATCAAGAGATTCATTAATTAATTAATTAATTAATTAATTA 2449  
 QY 1143 ATTAATTTCAATTTTAAAGCAATTTATTTTATTAATTAATTAATTAATTAATTAATTA 1202  
 DB 2450 CTGCTGATGATCTTTGAA-----TTTTATGAAGAAAGCGCAACTACTCTTAATTAATTT 2504

```

QY 1203 TTATAGTGGTATTTCTATCGTATGACCTTATATAGCAATATCTATTAATCTGAAAATT 1262
Db 2505 TTTTACCGCCATTAATATATGTTTCATTAACACTATATATATATCCAAAAATCTAG 2564
QY 1263 ATATGAGAAAAGAACAGGTTCACCCACAACAAAAAATATAGACATTTGATCTTTAA 1322
Db 2565 TGTTTTGGAAATCACAATGTATCTATATATATATATATATATATATATATATAT 2624
QY 1323 AGTTCAATGTATAGTATAGCAATCACTCTGTTTCCCTATATCAACCACTTAT 1382
Db 2625 TTATATTTTTTATTAATATGATATAGCTTATATATATATATATATATATATAT 2684
QY 1383 AATTATCAATTTGACTTTTATTTAATGCTCATCTAACACACACTCAAAATTTAGC 1442
Db 2685 TATTAGTAAATGATTTTTTATATATATATATATATATATATATATATATATAT 2744
QY 1443 AGGAGGCT-----TTATCTAATATCAAAAACAACCTTTTTTCAATTTCTAGAA 1496
Db 2745 AGCAGGATCTGGGCAAAATTAATCTTATGATGTAAATATATATATATATATATAT 2804
QY 1497 AAAAGACTGCAATCTAGTATATGATCCAGTTGTTCAACAACTTATATATATGTC 1556
Db 2805 TAAACGAAGAGATCAAGAAACCTTACCTTTTCCAAATATATATATATATATATG 2864
QY 1557 TATTTTATCCATTTTTCATTTATTTACTTATTTCTATGATGATGATTAACGTT 1616
Db 2865 TATTTTATC-----ATTATATAAAGTCTTATGATCCCTGCAACATAT 2907
QY 1617 ATTGATATCAGGTGTAT--TTAGATGACACACAGTATGTTGATATATATATATAT 1674
Db 2908 AAAATCTAGGTATATAGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 2967
QY 1675 TCAGATAAATATATTAATATATATATATATATATATATATATATATATATATAT 1734
Db 2968 TATACACTTATTAATCACTCAATATCCAGTGTAAAGGAATTCAGTGGAGCTGCT 3027
QY 1735 AAGGTATTTGAGGCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 1794
Db 3028 AAGGTGTTCAAGGCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 3083
QY 1795 CGTTAGAAATTTACATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 1854
Db 3084 --TTTCAAAATTTACATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 3141
QY 1855 TATGCTCAAAATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1914
Db 3142 TATGCTCAAAATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 3199
QY 1915 ATAGGATATCACTCAAGACTCAACAACTTTTCTGTTGTTGTTGTTGTTGTTG 1974
Db 3200 -TAGCAGAACTGGGATGCTCACTCAACCCACTTTTCTGTTGTTGTTGTTGTTG 3258
QY 1975 CAATACGAGATTTGGGTTATTTTCCAAATTTCCAAAGTATAGTATATATATAT 2034
Db 3259 AAATATTAAGATTTTCACTTATGAAATTTTCTAAAGGATGAAATTTGCTCCAAAT 3318
QY 2035 AACATACCATTTATATTTATATGTCAGATGAT---CAATTCATTTTATATGAT 2091
Db 3319 AACATATCTCTGTTGTTTATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 3378
QY 2092 AAAATGTAATTTATATCAATTTCTCTATGCAACAAATATAGAAAAAATTTA 2151
Db 3379 AAAATGTAATTTCTGCAATTTACTGTTCTATATAGAGAGATAGAGAAACAA 3438
QY 2152 GAAACTATCAAAACAAAATATATATATTTTCCAAATCACTACAAAAA 2200
Db 3439 GAAACAGTACAAACAATATATATATATATATATATATATATATATATATAT 3487

```

RESULT 7  
AAQ14669  
ID AAQ14669 standard; DNA; 3543 BP.

```

XX AAQ14669;
AC 27-AUG-2003 (revised)
XX 25-MAR-2003 (revised)
DT 04-FEB-1992 (first entry)
XX Dipteran active toxin gene.
DE Insecticide; B.t.; crystal; delta endotoxin; cryIIA; ss.
XX Bacillus thuringiensis serovar morrisoni.
XX Key Location/Qualifiers
FT 1..3543
FT CDS /*tag= a
XX EP457498-A.
XX 21-NOV-1991.
XX 09-MAY-1991; 91BP-00304180.
XX 15-MAY-1990; 90US-00524255.
XX 01-OCT-1990; 90US-00590903.
XX (MYCO ) MYCOGEN CORP.
XX Sick AJ;
XX WPI; 1991-341902/47.
XX P-PSDB; AAR14373.
XX Bacillus thuringiensis genes encoding diptera-active toxins - and
XX transformed microbes used to control insects in various environments.
XX Claim 1; Page 10; 20pp; English.
XX The sequence was obt. from plasmid pMYC1625 which was isolated from a
XX genomic library prep. from DNA from B.t. PS71M3 [from B.t. PS71M3-69
XX (NRRL B-18515)]. It is related to the cryIIA family of genes, the 140 kD
XX endotoxin gene and the type II gene from B.t. var. israeliensis. The gene
XX encodes a 130 kD protein. Microorganisms transformed with the DNA may be
XX administered to dipteran insects or their environments, the expressed
XX toxins acting as an insecticide. See also AAQ14670-Q14672. (Updated on 25
XX -MAR-2003 to correct PA field.) (Updated on 27-AUG-2003 to correct OS
XX field.)
SQ Sequence 3543 BP; 1278 A; 567 C; 612 G; 1086 T; 0 U; 0 Other;
Query Match 15.4%; Score 340.6; DB 2; Length 3543;
Best Local Similarity 53.1%; Pred. No. 6.5e-50;
Matches 1170; Conservative 0; Mismatches 934; Indels 99; Gaps 17;
QY 10 AATPACATATATACGAATATGAATTTGATTTGCGATACCTTATATTTTCCGAC 69
Db 4 AATCCTTATCAAAATATATATATATATATATATATATATATATATATATAT 63
QY 70 AGAAACGATATGTTTATGATCCCTTACACAAATATCAATCAACCTTACAAAC 129
Db 64 ATATCTATATATATATATATATATATATATATATATATATATATATATATAT 123
QY 130 ACAATTTCAAGAGTGGCTCAATATGTTCAAGGATATACAAATATGTTGATTTTC 189
Db 124 ACAATTTATTAAGTTGGCTCAATATGTTCAACAGATATGAGATATGTTGAGATTTT 183
QY 190 GAGACATTTGCTAGTGTGATATCAATTTGCTGCAATTTGTCAGGTATATTTTCCGT 249
Db 184 GAAACTTTTATGAT-----AGTGTGAATCTCAGTGTCTATATATATATATAT 234
QY 250 ACTGTGTTAGCCGATATAGTGGGCTCATTTATATCCGGAACCGATATATATAT 309
Db 235 ACCGTACTGACTGGT-----TTGGGTTTCAACAAACCTTTAGGA 273

```

QY 310 GCTATATATATCTTTGGTACCCCTAATCACTGCTTTTGGCCCGCGGAGAACAGAC 369  
DB 274 CTGTCTTAATATAGTTTGGTACATTAATACAGAGTTCTTTTCCAGCCCAACCAATCT 333  
QY 370 AAAACAGTATGACACAAATTTATTAATATGGAGAAATTTTGTGATACCCGTTTACA 429  
DB 334 AACAC---ATGAGAGTACTTTATTAACCAAACTAAATATTTATTAATAAATAATAGCA 390  
QY 430 -GAAAGCATTAACAGCTTAAGTTTCAAACTTTAGAGATTAGCAAAATTTTCAAG 488  
DB 391 TCAACATATATATATGCTATATTAATTAACAGGTCGTTTATGTTATACACACT 450  
QY 489 CTATATATCAGAGTTAGATGATTGGAGAAATTTAAAGATACACAGCTCTGATTAAC 548  
DB 451 TATCATATACCTTTAAACATGGAGAAATATCCAAACCCCAAAATCTCAGATATGA 510  
QY 549 ACCATCATCAGATTAACAACAAGCTGCTTGACTCTTAATAATACGATTTGAGAAATGTTCA 608  
DB 511 AGGA----CACAAATCCAGCTAGTTTATTAACATTTTCAAAATGTCATTTCCAGACTTGT 566  
QY 609 CAATGATTTTATTCAGAAATACCTGTTTCCAACTGTAATCTTAATAAGCTATTAAT 668  
DB 567 AAACCTTGTCTCTTAATCTAGTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 620  
QY 669 ACCATTTATGCGCAAGCTGCTAATTTTCAATTTTAAATTTTATTAACAACAGCTGTAAT 728  
DB 621 ATCTATGTTATGACAAAGCAGCAAACTTACATCTGATGATTTAAATCAAGCGCTCAAT 680  
QY 729 GCGTATGATGAAATGACATATATACATCTTCACAAAATGAACTTAATGCTGAAACATC 788  
DB 681 TGAACCGTATTTAAATAAACAATGCAATTCATTTATTTAGAGCTTGGCC---AACGC 737  
QY 789 AGATGACTATTAATTAACCTTTTAAAGAAATATACCTTAATATAGTAACTATTTGGCAA 848  
DB 738 AATTAATTAATTAATCAATTAATGACTTAAGCTTAATGAAGTTTACACTAATTTATTTGTAAC 797  
QY 849 TACCTATAGAACAGAGCTTAATAATCTTTAGAGACCAACCAATTAATGAATGAGATAT 908  
DB 798 AACTTATTAATAAAGATTTAAATTTTAAACGACCGCTGATGATTAATCTTGAAGAAA 857  
QY 909 TAAATGATATGAGATATATAGCACTTATCTGATTTAGATACATCTCTCAATTTTCTT 968  
DB 858 TATAATATGGAACATTA-----CAATGCTATGGAACAAATAATGACTAGCTGAT 911  
QY 969 ATATGATATTAATAAAGATATAGATTTCAATAGAGAAATGAAGTAAAGCATTTAAGAA 1028  
DB 912 AGATCTTGTGCACTCTTCTTAATTAATGATGATTAATATTCATTAATGAGTGCATATC 971  
QY 1029 TGAATCTCAAGAGAAATTTATTAACAATGAATTAATTTGATGCTTCTCAACTTAC 1088  
DB 972 TGAATCTTACTGAGAAATTTAT--CAAGTACTTAATCTTGAAGAAAGCCCTATTAATAT 1029  
QY 1089 AGTTCAACCCATCTAGCTACAGATGAATTAATTTAAACAGCTGCAAGTTTAAATTA 1148  
DB 1030 TATGACTTTCAATATCAAGAGATTTCACTTACACGTAAC--GCATTTATTTACTGGC 1087  
QY 1149 TTGATTTTGAACAATTTATTTTATTAACAGAAATACAAATTTTGGGAATGCTTACT 1208  
DB 1088 TTGATTTCTTGA-----TTTATTAAGAAAGCGAAATCTCTTAATTAATTTTTCAC 1142  
QY 1209 TGGTATTTCTAATCGATGACACTACTTAATGCAATCTAATTAATGAACTTTATATG 1268  
DB 1143 CAGCATTAATTAATATGTTTCACTTACACACTGATTAATATATATCCAAATCTAGCTTT 1202  
QY 1269 AGAAAGAACAGGTTCAACCAACAATAAATTAAGACATTTGAATCTTATTAAGTTTC 1328  
DB 1203 TGAATATCAATGTAATGATTAATTAATCTCTGCTTGGCAACAATAATTTATAT 1262  
QY 1329 AATGTAATGATTAAGCAATACCTGCTTCCCTATTTGAACGACCTTATATATTA 1388  
DB 1263 TTTTATTAATATGCTATGATTAATTAATTAATTAATTAATTAATTAATTAATTA 1322

QY 1389 TCAATTAATTAATTAATTAATTAATGCTATTAACAACAACACTCAAAATTTACAGCAGAG 1448  
DB 1323 TAAATGATTTTATTTATTAATTAATTAATGATTAAGTAACTTTTGGAGAAAGACTTACAGCAG 1382  
QY 1449 GTCT-----TATCTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1502  
DB 1383 ATCTGGCAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1442  
QY 1503 CTGCAATCTAGTTATTTGATTCAGAGTTTGTCAACCAACTTAATTAATTAATTAATTAATTAAT 1562  
DB 1443 AAGAGAAATTAAGAAATCCCTACCTTTTCCAAATATATATATATATATATATATATATAT 1502  
QY 1563 ATCCCATTTTCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1622  
DB 1503 ATC-----ATTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1545  
QY 1623 TACAGTGTATTTAGATGACACAGTATGTTGATTAATTAATTAATTAATTAATTAATTAATTAAT 1680  
DB 1546 CAAGTATTAAGTTTGTGCTGACACACTCTAGTGTATCTTAATAAATAATTAATTAATTAATTA 1605  
QY 1681 AAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1740  
DB 1606 CATTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1665  
QY 1741 ATTAAGAGACCTGATCATTAAGAGAACTGTTTATTAATTAATTAATTAATTAATTAATTAAT 1800  
DB 1666 GTTCAAGACCTGGTCAATTAAGAGAGGATTTAATTAATTAATTAATTAATTAATTAATTAATTA 1719  
QY 1801 GAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1860  
DB 1720 AAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1779  
QY 1861 ACAAATGCTGTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1920  
DB 1780 TCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1836  
QY 1921 ATACCACTCAACGATCAACAACAATTTTCTGCTAACAATAATTAATTAATTAATTAATTAAT 1980  
DB 1837 GAACTGGATTAAGGACCTCAACCCCACTTTTCTGGTACAGATTAATTAATTAATTAATTAAT 1896  
QY 1981 GGAATTTTGGGATTTTCAATTTTCAAGTACAGTACATTAATTAATTAATTAATTAATTAATTA 2040  
DB 1897 AAAGATTTTCAATTTTCAATTTTCAAGTACAGTACATTAATTAATTAATTAATTAATTAAT 1956  
QY 2041 CCATTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2097  
DB 1957 TCTCTGTTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2016  
QY 2098 GAATTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2157  
DB 2017 GAATTTTCTGCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2076  
QY 2158 ATCCAAACAATAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2200  
DB 2077 GTACAAACAATAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2119

RESULT 8  
AA081178  
ID AA081178 standard; DNA; 3543 BP.  
XX  
AC AA081178;  
XX  
DT 25-MAR-2003 (revised)  
XX  
DT 12-AUG-1995 (first entry)  
XX  
DE B.c. toxin P871M3 gene.  
XX  
KM Delta-endotoxin; crystal protein; biological control agent; Calliphoridae;  
XX  
KM screw-worm; sheep blowfly; Lucilia; Phormia; Calliphora; insecticide;  
XX  
XX pesticide; B.c.; ss.  
OS Bacillus thuringiensis.

XX W09502694-A2.  
 XX 26-JAN-1995.  
 XX 13-JUL-1994; 94MO-US007902.  
 XX 15-JUL-1993; 93US-00093199.  
 XX (MYCO ) MYCOGEN CORP.  
 XX Hickle LA, Payne J;  
 XX WPI, 1995-067338/09.  
 XX P-P8DB; AAR63078.  
 XX Method for controlling *Caliphoridae* pests - specifically *utillises*  
 PT *Bacillus thuringiensis* isolates or toxins.  
 XX Dielosure; Page 36-38; 50pp; English.  
 CC A library was constructed from *Bacillus thuringiensis* P871M3 total  
 CC cellular DNA in lambda Gem-11. Plasmid pMYC1625, selected in *Bac*herichia  
 CC coli, contained a 8.0 kb insert expressing a beta-endotoxin gene. This  
 CC was sequenced (AAQ81178). A cured, acrycetaliferous B.c. host carrying  
 CC pMYC1625 produced a 130 kDa crystal protein (AAR63078) toxic to  
 CC caliphorids. (Updated on 25-MAR-2003 to correct PN field.)  
 XX  
 XX Sequence 3543 BP; 1278 A; 566 C; 612 G; 1087 T; 0 U; 0 Other;  
 SQ  
 Query Match 15.4%; Score 340.6; DB 2; Length 3543;  
 Beest Local Similarity 53.1%; Fred. No. 6.5e-50;  
 Matches 1170; Conservative 0; Mismatches 934; Indels 99; Gaps 17;

QY 10 AATAACGATTAATGCAATATGAAATTAATGATTCGACATCTGACCTTAATTTCCGAC 69  
 DB 4 AATCTTATCAAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAAT 63  
 QY 70 AGAAACGATTAATGATTTAGATACCTTACACAAATTAATCCAAATCAATTAACAAAC 129  
 DB 64 ATATCTAATTAATTAATTAACAAATATCAATAGAAATAGTCCAAACAAATTAATTAACAAAGT 123  
 QY 130 ACAATTTGCAAGAGTGGCTCAATATGCTCAAGGAAATACAAATATGATGATTAATTC 189  
 DB 124 ACAATTAATTAAGATGCTCAATATGCTCAAGGAAATGCAAGATGATGATGATGATTT 183  
 QY 190 GAGACATTTGCTAGTGTGATACAAATGCTGCAAGTATGCAAGGATTAATTCGGT 249  
 DB 184 GAACTTTATGAT-----AGTGTGAACCTGAGTGTCTATCTATCTATGATGATGAT 234  
 QY 250 ACTCTGTAGCGGTATAGTGGGCTCACTTATATCCGACCGATAGAAATATAGT 309  
 DB 235 ACCGTACTGACTGGT-----TTCCGGTTCAACAACCTTATAGGA 273  
 QY 310 GCTAATTAATATCTTTTGTGACCTTAATCACTGCTTTTGGCCCGGGAACCAAGAC 369  
 DB 274 CTTCCTTAATAGGTTTGTGATCAATTAATACAGTTCTTTTTCAGCCCAAGACCAATCT 333  
 QY 370 AAAACAGTATGACACAACTTAATTAATGAGAGAAATTTTGTGATACACGTTTAA 429  
 DB 334 AACAC--ATGAGAGTACTTTTAACACAACTAAATTAATTAATTAATTAATTAATTAATTA 390  
 QY 430 -GAAAGCATTAAGAGCTTAAGTTTAACTTTAGAGAGATTTAGACAAATATTAACAAG 488  
 DB 391 TCAACATATATAGTAATGCTAATTAATTAATTAACAGTGTGTTAATGTTATACGACCT 450  
 QY 489 CTATTAATACAGATTTAGATGATGAGAAATTAATTAATTAATTAATTAATTAATTAATTA 548  
 DB 451 TATCATATATCACTTAATTAATGAGAGAAATTAATCAACCAAAATATCTACAGAGATGA 510  
 QY 549 ACCATCATACAGATTAACAAGAGCTGCTGACTCTTAATTAATTAATTAATTAATTAATTAAT 608  
 DB 511 AGGA-----CACAAATCAAGTATGTTCAATTCACATTTTCAAAATATGATTCAGAGCTTGT 566

QY 609 CAATGATTTTATGAGAAATACCTGTTTCAACTTGAACTTAATTAACGCTATTAAT 668  
 DB 567 AAACCTTGCTCTCTTAATCTAGTATGCT-----GATTAATTAACATTAATTAAT 620  
 QY 669 ACCATTTATGCGGACAGCTGTAATTTTCAATTAATTAATTAATTAATTAATTAATTAATTA 728  
 DB 621 ATCTAGTTATGACAAACAGCAACCTTAATGATGATGATGATGATGATGATGATGATGATGAT 680  
 QY 729 GAGTATGATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 788  
 DB 681 TGAAGCTATTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 737  
 QY 789 AGATGATTTATTAATCTTTTAAAGAAATTAATTAATTAATTAATTAATTAATTAATTAATTA 848  
 DB 738 AATGATTTATTAATCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 797  
 QY 849 TACCTATGAAAGAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 908  
 DB 798 AACTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 857  
 QY 909 TAATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 968  
 DB 858 TATTAATGGAACATTA-----CAATAGCTATGCAACAAATTAATTAATTAATTAATTAATTA 911  
 QY 969 ATATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1028  
 DB 912 AGATCTTGATGATCTCTTCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 971  
 QY 1029 TGAATCAAGAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1088  
 DB 972 TGAATCTTACGAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1029  
 QY 1089 AGTCAACCAATCTAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1148  
 DB 1030 TATGATTTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1087  
 QY 1149 TTCAATTTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1208  
 DB 1088 TTGATCTTTGAA-----TTTTATGAAAGGCAATCTAGTATTAATTTTTCAC 1142  
 QY 1209 TGGATTTCTAATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1268  
 DB 1143 CAGCATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1202  
 QY 1269 AGAAAGAAAGGTTCAACCAACAAACAAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1328  
 DB 1203 TGGAAATCAAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1282  
 QY 1329 AATTTAATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1388  
 DB 1263 TTTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1322  
 QY 1389 TCAATTTGAATTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1448  
 DB 1323 TAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1382  
 QY 1449 GTCT-----TTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1502  
 DB 1383 ATCTGGCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1442  
 QY 1503 CTGCAATCTAGTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1562  
 DB 1443 AAGAGAAATCAAGCAATCTTACCTTTTCCAAATTAATTAATTAATTAATTAATTAATTAATTA 1502  
 QY 1563 ATCCATTTTCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1622  
 DB 1503 ATC-----ATTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1545  
 QY 1623 TACAGGTGA--TTAGATGAGACAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1680  
 DB 1546 CAAGTGTATGCTTTGCTTGGACACACTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1605

QY 1661 AAAATTAATCAATGATCCAGCAATCAAGGTAACTATCTGTATACAACTTAAGTAA 1740  
 DB 1606 CATTTAACTACCAATTCACCTGTAAGGAATTCAGGGAGCTTCTTAAGGTT 1665  
 QY 1741 ATTGAAGACCTGCTGATACAGAGGAACTGGTTATTATTCAAAGTCAAGGGCGTTTA 1800  
 DB 1666 GTTCAAGACCTGGCTGATACAGAGGGGATTTAATTTGATTTCAAGATCA-----TTTC 1719  
 QY 1801 GAAATTAATCAATGAACTCTTAATCTACAACTATTTATTTGATTAAGCTTGATATGCT 1860  
 DB 1720 AAAATTAATCAATGAACTCTTAATCTACAACTATTTATTTGATTAAGATTTGATATGCT 1779  
 QY 1861 ACAAATGCTGCTGAATACTCTTCTATATATCTCTTACAAATCAAGAGTAATAGGA 1920  
 DB 1780 TCAATTAAGAGGCAATATACAGAGCTGTTATTAATCTTAAGATCCAGGGG---TAGCA 1836  
 QY 1921 ATACCACTCAACGACTCAACAACTTTTCTGGTACAAATTAATTAATTAACAATAC 1980  
 DB 1837 GAACCTGGATGAGCACTCAACCCACTTTTCTGGTACAGATTAACGAATTTAAATAT 1896  
 QY 1981 GGAGATTTGGGTATTTCCAAATTTTCCAAAGTACAGTAACCTTTAAATGAACAATA 2040  
 DB 1897 AAAGATTTTCACTACTTGAATTTTCTACAGAGTGAATTTGCTCCAAATCAAAACATA 1956  
 QY 2041 CCATTTATATTTAATCGTGCAGATGAT---CAAAATTCATTTTAATCATTTGATTAAT 2097  
 DB 1957 TCTCTTGTTTAATCGTTGAGATGATTAATCAACACAGATGATCTTAATGATTAAT 2016  
 QY 2098 GAATTTATCAATTAATCTCTCTATGCACCAAATTAAGAAAAAATAAATTAAGAACT 2157  
 DB 2017 GAATTTCTGCCAATTAATCTCTCTATTAAGAGAGATAGAGAAACAAAAATTAAGAAACA 2076  
 QY 2158 ATCCAAACAAATTAATCAATTTTTCAAATCTATACAAAAA 2200  
 DB 2077 GTACAAACAAATTAATCAATTTATGCAAAATCTATAAAAA 2119

## RESULT 9

AAQ14670  
ID AAQ14670 standard; DNA; 2061 BP.

AC AAQ14670;  
 XX  
 DT 27-AUG-2003 (revised)  
 DT 25-MAR-2003 (revised)  
 DT 04-FEB-1992 (first entry)  
 XX  
 DE Dipteran active toxin gene.  
 XX  
 KW Insecticide; B.t; crystal; delta endotoxin; cryIIVC; ss.  
 XX  
 OS *Bacillus thuringiensis serovar morrisoni*.  
 XX  
 Key Location/Qualifiers  
 FT CDS 1..2028  
 FT /\*tag= a  
 XX  
 EN EP457498-A.  
 XX  
 PD 21-NOV-1991.  
 XX  
 PF 09-MAY-1991; 91EP-00304180.  
 XX  
 PR 15-MAY-1990; 90US-00524255.  
 PR 01-OCT-1990; 90US-00590903.  
 XX  
 PA (MYCO ) MYCOGEN CORP.  
 XX  
 PI Sick AJ;  
 XX  
 DR WPI; 1991-341902/47.  
 DR P-PSDB; AARI4374.  
 XX

PT *Bacillus thuringiensis* genes encoding diptera-active toxins - and  
 PT transformed microbes used to control insects in various environments.  
 XX  
 PS Claim 1; Page 15; 20pp; English.

The sequence was obtd. from plasmid pMYC1636 which was isolated from a  
 CC genomic library prepd. from DNA from B.t. PS71M3 [from B.t. PS71M3-69  
 CC (NRRL B-1815)]. It is related to the cryIIVC from B.t. var. *israelensis*.  
 CC The gene encodes a 77 kD protein. Microorganisms transformed with the DNA  
 CC may be administered to dipteran insects or their environments, the  
 CC expressed toxins acting as an insecticide. See also AAQ14669-Q14672.  
 CC (Updated on 25-MAR-2003 to correct PA field.) (Updated on 27-AUG-2003 to  
 CC correct OS field.)  
 XX  
 SQ Sequence 2061 BP; 779 A; 331 C; 319 G; 632 T; 0 U; 0 Other;

Query Match 8.5%; Score 188.2; DB 2; Length 2061;  
 Best Local Similarity 50.0%; Pred. No. 1.9e-23;  
 Matches 1107; Conservative 0; Mismatches 923; Indels 185; Gaps 18;

QY 1 ATGAATCAAAATTAAGATTAATACGAATATGAAATTAATGATGCAATACCTCACTTAT 60  
 DB 1 ATGAATCAATTAATCAAAATTAAGATTAATGAAATTAATGATGCTC-----CATCAAT 54  
 QY 61 TTTCGAAACAGAAACAGTAATGATTTAGATACCTTACACAAATTAATCAAAATCAACA 120  
 DB 55 GGTTTTACAGAGCTTAATTAATTAATTTATGATATCATTAAGCAAAATTAAGCAAAATCAACA 114  
 QY 121 TTACAAACACAAATTAACAAAGAGTGGCTCAATATGTCGCAAGGAAATACCAATATGCT 180  
 DB 115 CTGAAAAACAGAAATTAACAAAGATTTGGCTCAATGTCGTCGCAAGATTAATCAAAATATGCG 174  
 QY 181 GATAATTTGAGACATTTGCTAGTCCGATATCAATTTGCTGAGTAGAGCAAGTACTATT 240  
 DB 175 AATATGCGGGAAATTTTGTATGTAACCTATGTTGAGATGAGCAAGATTAATTT 234  
 QY 241 GTATCCGCTACTCTGTTAGCCGATATAGTGGGCTCACTTCAATTCGGAACCGATAGGA 300  
 DB 235 GTAGAGGAACCTATATGTTAGAGCTTTTGCTGCCCT----- 270  
 QY 301 ATAAATGCTGCTATTAATTAATATCTTTTGGTACCTTAATCACTGCTTTTGGCCCGGGA 360  
 DB 271 GTCTTAGCTGAGGATTAATTAATCTTTTGGGACTTGTGTGCGAATCTTTTGGCAAGATCT 330  
 QY 361 GAACAAAGCAAAACAGTATGAGCACAAATTTAATAAAGGAGAAATTTTGTGATACA 420  
 DB 331 GACCCTG---CAAAATGTTGGCAGGATTTGTTAAACATCGAGGAGGCTTATACAGAA 387  
 QY 421 CCGTTTACAGAAAGCATTAACAGCTAAAGTTTACAACTTTAGAGAGATTTAGACAAATA 480  
 DB 388 ATAGATTAATAACATTAATTAATGATGATCTA-----CTATCGTAACCTTATTAATAAAT 441  
 QY 481 TTACAAAGCTATTAATTAACAGCAATTAAGATGAGAAAAATTAATAAGACTACAGCTCT 540  
 DB 442 CAACCTGATTAATTAATTAACAGAAATTTTTCGATTAATGAGGAGCAGACAGTACACCTAT 501  
 QY 541 GGATTAACAGCATCATGACATTAACAACAGCTGCTGACTTAATAATTCGATTTAG 600  
 DB 502 GCTAAGACAGTA-CATGATCTCTTACTACTAGAACCTTAATTAAGTAAGATTAAGA 560  
 QY 601 AATGTTCAATGATTTTATTCGAGAAATATCTGTTTCAACTTGAACCTTATTAAG 660  
 DB 561 TATGTTAAAA-----ATAATGCTAGCTATGGAATA 591  
 QY 661 CTATTTACTACTTATTAATGAGCAAGCTGCTAATTTTCAATTTAATTAATTAACAACAGGT 720  
 DB 592 CCAACACTCCCTGATATGACAAATATGCTACTGGCACTGGAATTTAATTAACATCT 651  
 QY 721 GCTGAATGCTGATGATGAGATGCAATATTAATCTTCAAAATTAAGCAATGCACTTAATCT 780  
 DB 652 GCTACCTATTAACAAATTAATGCTGCAAAATCAAGGAT-----AAATCAAGTACTTTC 705  
 QY 781 GGAACATCAGATGACTATTAATTAATTTTAAAGAAATATATCTTAATATATAGTAATCT 840



1 ATGATCAAAATPACGATATATACGAATATGAAATTAATGATTCGATACCTCACCATT 60  
1 ATGAAATCCATATCAAAATPAGAAATGAATGAATATCAATGCTC-----CATCCAAAT 54  
61 TTTCGGAAGAAACAGTAATGATTTCTATATACCTTACCAAAATATCCAAATGACCA 120  
55 GGTTTTAGCAAGCTATATACATATCTTAATCTAATATCCATAGCAAAATAGCAAAATCAACA 114  
121 TTACAAAACAAATTTACAAAGTGGCTCAATATGTGCAAGGAAATACAAATATGCT 180  
115 CTGAAAACACGAATTTACAAAGTTGGCTCAATGTGTGCAAGATATACAAATATGCG 174  
181 GATTAATTTGAGACATTTGCTAGTGTGATACAAATGTGCAAGTTAGTGCAGTACTAT 240  
175 AATTAATGCGGGGAATTTTGTAGTTCTGAACATATGTTGAGGATTAAGTCAAGTATAT 234  
241 GATTCGGATCTCTGTAGCCGGATAGTGGGCTCACTTATATCCGACCGATAGGA 300  
235 GTATAGGAACCTATGTAGGAGCTTTTGCTGCCCT----- 270  
301 AATATAGGCTATATATATATCTTTTGGTACCCCTAATGACTGTCTTTGGCCGGGGA 360  
271 GTCTTAGCTGCGATATATATCTTTTGGGACTTGTGCTCCAAATCTTTTGGCAAGGATCT 330  
361 GAACAAGACAAACAGTATGACACAAATTTATTAATGGAAGAAATTTTGTGATACA 420  
331 GACCCGTG---CAATGTTGGGAGGATTTGTTAAACATCGAGGAAGGCCCTATACAGAA 387  
421 CCGTTACGAAAGCTTAAACAGCTTAACAGTTTAACTTTAGAGGATTTAGCAAAAT 480  
388 ATATATTAATAACATATATATATATGATCTTACCTT-----CTATCGTAAACCTTATTAATAAT 441  
481 TTACAAAGCTATATATATATATCTTTTGGTACCCCTAATGACTGTCTTTGGCCGGGGA 540  
442 CAATCTGATTAATATATCAAGAAATTTTTCGATTAATGGAAGCCAGCAGTACACAGCTAAT 501  
541 GGATTAACCAACATCATCAGCATTTACAACAAGCTGCTTACTCTTAAATAGCAATTTGAG 600  
502 GCTAAAGGAGTA-CATGATCTCTTATCTACCTTAGAACCTATATATGATTAAGATTTTAA 560  
601 AATGTTACAGATGATTTTATTCGAGAAATACCTGGTTTCCAACTTGAAACTTATTAACG 660  
561 TATGTATTAATA-----ATATGCTATGCTATGCAATA 591  
661 CTATTACTACCTATTTATGCGCAAGCTGCTAATTTTCACTTAATTTTATTAACAAGGT 720  
592 CCAACATCTCCGCAATATGCAAAATAGCTACTGCGACTTAATTTATTTAAACAAGCT 651  
721 GCTGAATTTGGCTGATGAAATGGAATGCAATATCACTTCAAAATTTGAACCTATATGCT 780  
652 GTTACCTTATTAATATATATGCTGCAAAATCAAGGAT-----AATTCGAAGTACTTTC 705  
781 GGAACATCAGATGACTATTTATTAACCTTTTAAAGAAATATATCTTAATATATATATAT 840  
706 AATTCATCTAATTTACTATCAGGGCTATTTTAAACGTAATATCAAGAAATATCTGACTAT 765  
841 TGTGCAATATCTATATAGAACAGAGCTTAATAATCTTTAGAGCAAGCAAAATATGAAATG 900  
766 TGTATATCAAAACGTACAAATGCAAGACCTAATGATTTAGAACTAATATATCAACCAATG 825  
901 AGTATATTTAATGACTATGGAAGATATATGACCACTTACTGTATATGATCCATCTCTCA 960  
826 AATATGATATATATCTTACCGTTTAAAGATGACTTACCTGTGTATGATTTTATGCTAAT 885  
961 TTTTCTTATATGATATTAATAAGATATAGATTTCAATAGAGGAATATAGATTAAGGC 1020  
886 TTTCCAAATATATGACCCCAAAAATATCC-----AATAGGA 921  
1021 AATTAAGATGACTGACAGAGAAATTTATACACTGA-AATTAATTTTATGATGCTTCC 1079  
922 GTTAATATCTGAACCTTACAGAGAAATTTATAGCAATGTTTAAATTCAGATCATTTTGAACC 981  
1080 TCAACTTAGAGTTCAACCAATCTAGCTAGATGAGATTAATTTTAAACGTCGAAGTTT 1139

982 ATTAACAGAACTAGAAAATGAGTATTACTGAATCCCTACATATTTTATCTTGATTAACCA 1041  
1140 TAAATTTATCTATTTTATGAAACAATTTATTTTATACAGAAAATPACAAATTTGGGGA 1199  
1042 GGGCGTTTATACAGAAATTTCTGAGCAATTTCTTGATCTCTTATGATATTTTCTTTT 1101  
1200 TCGTTAGTTGGTATTTCTAATCGTATGACCTTATATAGCAATATATATATGTAAC 1259  
1102 ACAGGTAACAG-ATGGCCTTTACATATATGATGATGCGCAATATATCTGGGAGC 1160  
1260 TTTATATGAGAAAGAACAGGTTCAACCACAAAGAAACATATAGACATTTGAATCTTA 1319  
1161 GGTTCATGACATATATTTTCTCAAGACATATCCAAA-----GTATTTCC 1205  
1320 TAAATTTCAATATGATACGATATAGCAATACCTCCGTTTCCCTATTTCAACCAACTT 1379  
1206 TTTTATATGAAACAAACCTATATGATTAAGGTGAAATTTGTCAGACATATAGAGTACTAGA 1265  
1380 TATATATATCAAAATTTGAACCTTATTTAATGCTCATCTAACAACACTCAAAATATTC 1439  
1266 TATATATATGAAATGATATTTTTCGAAATAGC-----AGTGAAGTATTTGCAATATTC 1319  
1440 AGCAGAGGGTCTTTATCTAATTTATCAAAACAACTTTTTCATATTTCTAGAAAAA 1499  
1320 ATCCAAATCAACAATAGAAAATTAATTAAGAACTGATCTTATATGATTTCAAAACA 1379  
1500 AGACTGCAATCTAGTATTTGATCCAGGTTGTTCAACCAACTTATATATATATAGCTAT 1559  
1380 AACATG-----GAAAATTAAGAAATATGTCATATC 1409  
1560 TTTATCCATTTTTCATTTTACTTATTTCTATGATGATTTGATTTACGCTCAAAATAT 1619  
1410 TCTATCTGATATTAATACTGATATATAT-----ATATTTTCAAGTATAGGAAAG 1460  
1620 AGATACAGGTGATATGATGAGACACACAGTATGTTATATGATATATGCAATATCAGA 1679  
1461 AAGAAAGATTCATTTAGTTGACACATATAGTGTATTTCCAAATATCAATATGATTT 1520  
1680 TAAATTAATTTCAATGATCCCAAGCAATCAAGGTAAACATCTGATCAAAATCTTAAAGT 1739  
1521 AGATTAACATCAACCAATTCACAGCTTAAGCTTAAAGCTTGAAGTATGATTTCAAAAT 1580  
1740 AATTAAGACCTGGTCAATACAGAGAACTTGGTTTATTTTACAAGTCAAGGCGTTT 1799  
1581 TGTGAAGGTCTGCTCACAAGGTGAGACTTGTATTTCTTAAGATATAGATGATTT 1640  
1641 TAGAGTTAGATTTTAAATAATGTTTCTGACAAATATCAAGTACGTA---TTGCTTATGC 1697  
1800 AGAAATTAATGATGAACTCCTAATTTCTACAAATCTTATTTCACTTAGACTTGCATATGCT 1859  
1641 TAGAGTTAGATTTTAAATAATGTTTCTGACAAATATCAAGTACGTA---TTGCTTATGC 1697  
1860 TACAAATGCTGTGAATATCTTCTCTAATATATCTTCAATATCAAGAGTAAATAG 1919  
1698 TACTATATCTCCAAAGACAAAGCT--ATTCCTTAACCGGAATAGATATATAGTGTGAG 1755  
1920 AATACCACTCAAGCACTCAACAACCTTTTCTGCTCAATATTAATTAATTTTCAATA 1979  
1756 CTCCCTAGTACACTTCCCGCAAAAC-----CAATGCTACAGATTTTAACTA 1805  
1980 CGGAGATTTTGGTATTTTCAATTTTCAAGTACAGT-----ACATTAACCTTTTAA 2030  
1806 TGCAGATTTTGGATATGATATGATTTTCAAGAAACAGTTCCAATATTAACATTTTGAAGAGA 1865  
2031 TCGAAACATCACTTATATTTAATATGCTGCAATATCAAAATTTCAATTTTATATATGCA 2090  
1866 AGACACTTATTTAAGACTTAT 1925  
2091 TAAATTAATTTATACCAATTTACTCTCTATGACCAAAATAGAGAAAAACAATAAT 2150  
1926 CAAATTCGAATTTATTTCAATCACTCAATCTGTATATATATATATATATATATATATAT 1985  
2151 AGAAATCTATCAACAACAAATTAATATATATTTTCAAAATCATCAAAACACTT 2205

Db 1986 AGAAAAACACAGAAATAGTGAATTTATTTGTTAATTAACAAAGTTCTT 2040

RESULT 11  
AAN50525  
ID AAN50525 standard; DNA; 3756 BP.  
XX  
AC AAN50525;  
XX  
DT 24-OCT-2003 (revised)  
DT 25-MAR-2003 (revised)  
DT 23-OCT-1991 (first entry)  
XX  
DE Bacillus thuringiensis var. israelensis endotoxin insert in plasmid  
DS pS1367.  
XX  
KM Endotoxin; insecticide; ss.  
XX  
OS Bacillus thuringiensis serovar israelensis.  
XX  
PN EP153166-A.  
XX  
PD 28-AUG-1985.  
XX  
PF 15-FEB-1985; 85EP-00301017.  
XX  
PR 22-FEB-1984; 84US-00582506.  
PR 22-JAN-1985; 85US-00693556.  
XX  
PA (SYTR ) SYNTRO CORP.  
XX  
PI Walfield AM, Pollock TJ;  
DR WPI, 1985-211724/35.  
XX  
PT Polypeptide active against Diptera insects - prepd. from DNA sequence  
PT coding for Bti endotoxin using bacterial host.  
XX  
PS Disclosure; Page 15a-e; 27pp; English.  
XX  
CC The B. thuringiensis var. israelensis endotoxin insert in pS1367 is  
CC expressed in a bacterial host. The protein produced has insecticidal  
CC activity against dipteran insects. (Updated on 25-MAR-2003 to correct PA  
CC field.) (Updated on 24-OCT-2003 to standardise OS field)  
XX  
SQ Sequence 3756 BP; 1379 A; 566 C; 659 G; 1150 T; 0 U; 2 Other;

Query Match 8.1%; Score 178.2; DB 1; Length 3756;  
Best Local Similarity 50.0%; Pred. No. 1,1e-21;  
Matches 1108; Conservative 0; Mismatches 923; Indels 186; Gaps 19;

QY 1 ATGAATCAAAATACGATTAATACGAATATGAATTTATGATTCGATACCTCACTTAT 60  
DB 941 ATGATTCATATCAAAATATAGAAATGAAATATTCATTCATTC-----CATCCAT 994  
QY 61 TTTCCGAACAGAAACGATATGATTTCTAGATACCTTACACAAATATCCAAATACCA 120  
DB 995 GGTTTTACGAACTCTAATTAATCTATCTAGATATTCATTAGCAAAATAGCCAAATCAACCA 1054  
QY 121 TTACAAAACAAATTTACAAAGAGTGGCTCAATATGTGCAAGGAAATACAAATATAGT 180  
DB 1055 CTGAATAACAGCAATTTACAAAGATGGCTCAATGTGTGCAAGATTAATCAACAAATATG 1114  
QY 181 GATTAATTTGAGACATTTGCTAGTGTGCTGATACAAATTTGCTGACAGTATGACGATATT 240  
DB 1115 AATATATGGGGGAATTTTGTCTAGTTCTGAACTATTTGTTGAGGTATGAGATATTT 1174  
QY 241 GTATCCGCTGCTCTGTAGCCGGATATAGTGGGCTCACTTATATCCGACCGATAGGA 300  
DB 1175 GTAGTAGGAACATATGTTTAGAGCTTTGTGCCC-----CT 1210  
QY 301 ATTAATAGGCTATATATATCTTTTGTATCCCTAATCACTGTCTTTTGGCCCGGGGA 360

Db 1211 GTCTTACGCAAGTAAATATCTTTTGGGACTTTGTGCCGACTTTTGGCAAGATCT 1270  
QY 361 GAACAGACAAACAGTATGACACAAATTTATTAATGGAGAAATTTTGTGATACA 420  
DB 1271 GACCTG---CAATGTTTGGCAGAGATTGTTAAACATCGAGAGAGCCATATCAAGAA 1327  
QY 421 CCGTTAACAGAAAGCATTAACAGCTTAAAGTTACAACTTTAGAGAGATTTAGCAATA 480  
DB 1328 ATGATATTAACATTAATTAATGTAATCTT-----CTATCGTACACCTATTAATAAT 1381  
QY 481 TTACAAAGCTATTAATACAGCATTTAGATGATTTGAGAAATTTAAAGACTACAGCTCCT 540  
DB 1382 CACTGTATTAATATCAAGAAATTTTGTGATTAATGGAGCCAGACGATACAGCTAAT 1441  
QY 541 GGATTAACCAACATCAGCATTAACAACAGCTGCTTGAAGCTTTAAATATGATTTGAG 600  
DB 1442 GCTAAAGCAGTA-CATGATCTCTTACTTACCTTGAAGCTATATATGATTAAGATTTAGA 1500  
QY 601 AATGTTCACAATGATTTTATTCGAGAAATACCTGGTTCCAACTTGAACTTATTAACG 660  
DB 1501 TATGTTAAAA-----ATAATGCTATGCTATCGAATA 1531  
QY 661 CTATTACCTATTTATGCGCAAGCTGCTAATTTTCAATTTAATTTATTAACAACAGT 720  
DB 1532 CCACACTCCCTGCTATATGCAAAATAGCTTGTGCACTTGAAATTTATTAACATGCT 1591  
QY 721 GCTGAATGGCTGATGATGAAATGCAATATACATCTTCACAAAATGAACTTAATGCT 780  
DB 1592 GCTACCTATTAACAATATATGCTGCAAAATCAAGATTAATCCAGATCTTCAATTT-- 1649  
QY 781 GGAACTACAGATGCTATTTAATTTTAAAGAAATATACCTAATATATAGTACTAT 840  
DB 1650 ----CATCTAATTAATCTATCAGGGCTATTTAAAGTAAATATCAAGATATATCTACTAT 1705  
QY 841 TGTGCAATACCTATAGAACAGACTAATAATCTTATAGACGAAACCAATATGAATGG 900  
DB 1706 TGTATTAACAACGTAATGCAAGACTAATGATTTGAACTAATATCTTAACGCAACTGG 1785  
QY 901 AGTATATTTAATGACTATGCAAGATATATGACATTAATCTGATTTAGTATACCTCTCA 960  
DB 1766 AATATGATATATATCTTACCGTTTGAATGACTTAATCTGTTAGATTTATGCTATT 1825  
QY 961 TTTTCTTATATGATTAATTAAGATATATGAGATTCATATAGAGAAATAGAAATGAAGC 1020  
DB 1826 TTTCAATTAATGACCCAGAAATATATCC-----AATAGGA 1861  
QY 1021 ATTAAGATGAATCAGACAGAAATTTATACACTGAATTAATTTTGAATCGTCTTCT 1080  
DB 1862 GTTAAATCTGAACTTATCAGAGAAATTTATACGAATGTTAA----- 1902  
QY 1903 ----TTCAGTACATTTAGAACCATTAACAGAAATGGAATTAATCTAGAAATCTCT 1957  
QY 1141 AATTTATTTTCACTTTTATGAAACAATTTATTTTATACAGAAATATCAAAATTTGGGAAT 1200  
DB 1958 ACATTATTTTCTTGGAATTAACCAAGGCGCTTTTATACCAAGAAATTC--TCGAGACAT 2014  
QY 1201 CGTTAGTGTATTTCTAATCGTATGCACTCTATATAGCAATATCTAATCACTGAAT 1260  
DB 2015 CTGATCTTATATATATTTTCTTTTACAGTATACCAAGATGCGCTTTTACATATCTAT 2074  
QY 1261 TTATATGAGAGAAAGACAGCTTCAACCAACAAACAAATTAAGACATTTGAATCTTAT 1320  
DB 2075 GATGATCGCAACATTAATCTGGGAGCGGTTCAATGAATATTAATTTCTCAAGACATCTC 2134  
QY 1321 AAAGTTTCAATTTGAATGATGACATCACTCTGTTTCCCTATTTCAACGACATTT 1380  
DB 2135 AAAGTATTTCTTTTATGAAACAAACCTATGATTAAGTGAATTTGT--CAGACATA 2192  
QY 1381 ATTAATATCAAAATGAACTTATTTAATGGCTCATCTAACAACACACTCAATATTTCA 1440  
DB 2193 GAGAGTACTCAGATATATATATGAAATGATATTTTTCGAATAGCAGTGAATATTT- 2251



QY 1441 GGAGAGGCTCTTATCTAATTATCAAAAACAACTTTTTCATTCTTAGAAAAA 1500  
 DB 2252 -----CGATTTCATTCATTAACAAATAGAAAATATTAATTAAGAACT----- 2296  
 QY 1501 GACTGCATCTGATTATGATCCAGGTGCTTCCAACTTTAATTAATAGTCATATT 1560  
 DB 2297 -----GATTCTTATATGATTCCTAAAAACAACATGAAAAATGAAATATGCTACT 2350  
 QY 1561 TTATCCATTTTTCATTAATTACTTATTCCTATGATGATGATTAACGTACAAATATA 1620  
 DB 2351 CTATGCTATATATAAACTGATATATAT-----ATATTTTCAGTATGATGAGAAAGA 2401  
 QY 1621 GATACAGGTGATTTAGATGACACACAGTATGATGATATATATCAATATCAAT 1680  
 DB 2402 AGAAGGTTGATTTAGTTGACACATCTAGTGTGATTTCCAAATATCAATAGTTTA 2461  
 QY 1681 AAAATATATTAATGATGATCCAGCAATCAAAAGTAAATCTTGATTAAGGTA 1740  
 DB 2462 GATTAATATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAAT 2521  
 QY 1741 ATTGAAGACCTGTCATACAGAGAACTGTTTATTTTCAAAAGTCAAGGCTTTA 1800  
 DB 2522 GTGAAGGCTGCTGTCACAAAGTGAACCTTGATTAATCTTAAGATGATGATTTT 2581  
 QY 1801 GAAATTAATCAATGAAACTCTTAATCTTACACAACTTATTTAGATGATGATGCT 1860  
 DB 2582 AGAGTTAGTTTATTAATAAATGTTTCTCAACATATCAAGTACGTA---TTGCTATGCT 2638  
 QY 1861 ACAAAATGCTGGAATACTCTTCTATATATCTCTTACAAATCAAGAGTAAATGGA 1920  
 DB 2639 ACTAATGCTCAAAAGCAACAGTA--TTCTTAACGGAATAGTACTATAAGTCTGAGAC 2696  
 QY 1921 ATACCACTCAACGACTCAACAACTTTTCTGACAAATTAATTAATTAATTAATC 1980  
 DB 2697 TCCCTAGTACCACTTCCCGCAAAAC-----CCTAAATGCTACAGATTTAATCTAT 2746  
 QY 1981 GGAGATTTGGGATTTTCAATTTTCCAAATACAGT---ACATTAATCTTTAATGAAA 2036  
 DB 2747 GAGATTTGGATTTGATTAATCAATTTTCAGAAAGCTTTCAATTAATTAATTAATTA 2806  
 QY 2037 CATACATTTATATATTA-----ATCGTGAATGATGATCAATTTCAATTTTAAAT 2088  
 DB 2807 GACACTTTTATTAATGACCTTATATGATGATCAACCAATCATTCATTAATTAATTA 2866  
 QY 2088 GATTAATTTGATTTTATACATTTACTTCTCTATGACCAAAATGAGAAAAA 2148  
 DB 2867 GACAAATTTGATTTATTCATCTCAATCTGATTAATGATTAATGAGAAAGCAAAAT 2926  
 QY 2149 TTAGAACTATCAACAAATAATATACATTTTTCACAAATCATTCACAAACACTT 2205  
 DB 2927 ATAGAAAAACACAGAAATAGTAAATGATTAATTTGTTAATTAACAAAGTTCTT 2983  
 RESULT 12  
 AAD43974  
 ID AAD43974 strand: DNA; 4896 BP.  
 XX AAD43974;  
 AC  
 XX 13-DEC-2002 (first entry)  
 DT  
 XX Bacillus thuringiensis ssp. finitimus cry28a1 gene.  
 DE  
 XX Delta-endotoxin; cry26a1; cry28a1; insect-resistant plant; toxin;  
 KW transgenic host cell; insecticide; gene; ds.  
 XX  
 OS Bacillus thuringiensis.  
 XX  
 FH Key Location/Qualifiers  
 FT CDS 1129..4458  
 FT /tag= a  
 FT /product= "cry28a1 gene"

XX US2002038005-A1.  
 PN  
 XX 28-MAR-2002.  
 PD  
 XX 08-JAN-2001; 2001US-00756526.  
 PF  
 XX 07-JAN-2000; 2000US-0175158P.  
 PR  
 XX  
 PA (WOJC/) MOJICHOVSKA J A.  
 PA (LEWIT/) LEWITTIN E I.  
 PA (ZALU/) ZALUNIN I A.  
 PA (REV1/) REVINA L P.  
 PA (CHES/) CHESTUKHINA G G.  
 PI Mojicchowka JA, Lewitin EI, Zalunin IA, Revina LP;  
 PI Chestukhina GG;  
 DR WPI; 2002-403936/43.  
 DR P-PSDB; AAE26353.  
 PT Novel isolated delta-endotoxin nucleic acid molecules, cry26a1 and  
 PT cry28a1 isolated from Bacillus thuringiensis finitimus, that encodes  
 PT toxin active against insects, useful for controlling insects.  
 PS  
 XX Claim 1; Page 33-37; 42pp; English.  
 CC The invention relates to isolated delta-endotoxin nucleic acid molecules,  
 CC cry26a1 and cry28a1 isolated from Bacillus thuringiensis finitimus,  
 CC that encode a toxin that is active against insects. The invention is  
 CC useful for producing an insect-resistant plant, by introducing the  
 CC nucleic acid molecule into the plant, where the nucleic acid is  
 CC expressible in the plant in an effective amount to control an insect. The  
 CC invention is useful for producing a toxin that is active against insects  
 CC by obtaining the transgenic host cell and expressing the nucleic acid  
 CC molecule in the host cell, which results in the toxin that is active  
 CC against insects. The toxin is useful for controlling an insect by  
 CC delivering to the insect an effective amount of toxin. The invention is  
 CC useful for controlling insects. The toxin is useful for inhibiting the  
 CC ability of insect pest to survive, grow or reproduce, for limiting insect  
 CC -related damage or loss in crop plants, and to prophylactically treat  
 CC insect susceptible areas or plants to confer protection or resistance  
 CC against harmful insects. The present sequence is Bacillus thuringiensis  
 CC ssp. finitimus cry28a1 gene  
 XX  
 SQ Sequence 4896 BP; 1727 A; 757 C; 887 G; 1525 T; 0 U; 0 Other;  
 Query Match 7.4%; Score 163.8; DB 6; Length 4896;  
 Best Local Similarity 55.9%; Pred. No. 3.4e-19;  
 Matches 405; Conservative 0; Mismatches 272; Indels 48; Gaps 3;  
 QY 306 AGTGTCTATATATATCTTTTGTGATCACTGCTCTTTTGGCCGCGGAGAA 365  
 DB 1293 AGTGTGTACTTATATATATGGAACCTTGCTCCGTTCTTTGGCTGATCCAGAGA 1352  
 QY 366 AGA---CAAAACAGATGACACATTTATTAATTAATGGAAGAAATTTTGTGATACAC 422  
 DB 1353 AGATCAAAAAAATTTGTGTCATTTATGAAACAGAGAGACCTTTTAATCAAAAC 1412  
 QY 423 GTTAACAGAAAGCATTAACAGCTTAAGTTACAACTTTAGAGATTTAGCAATATTT 482  
 DB 1413 AATTTCTACAGCTGTAAGAAATATGATGCTCATCTAAATGTTTAAAGATGATTT 1472  
 QY 483 ACAAGCTATATATACAGCATTTGATGATGAGAAATTAATAAAGATCAACAGCTCTG 542  
 DB 1473 AACGTACATATGAAAGAGATTTAATGATGGAAGAAATCCMACTGCMAATACGCG 1532  
 QY 543 ATTACCATCATCATGAGCATTAACAACAGCTGCTGACTTAAATATAGCATTTGAGAA 602  
 DB 1533 ATTGTATACAGA-----GATTTGAAA 1556  
 QY 603 TGTTCACAAATGATTTTATTCGAGAAATACCTGGTTTCCAATCTTAATTAACGCT 662

Db 1557 CGCTCATTTCATTTTGTAGCAATATGCCACAACCTCCACGTTACCACTT 1616  
 Qy 663 ATTAACCTATTTATATGCGCAAGCTGCTAATTTTCAATTTAATTTATACAAAGATGC 722  
 Db 1617 ATTATTAAGTTGCTATACAGAAAGCTGCAGAAATTTACATTTGATTTATACATCAAGGTGT 1676  
 Qy 723 TGAATTTGGCTGATGAATGAATGCGAGATATATACATCTTTCACAAATTTGAACCTAATGCTGG 782  
 Db 1677 ACAATTCGCGATCATATGAAATGCGAGATCAACCAATTCACCAAT-----GTTGAA 1727  
 Qy 783 AACATCATGATGCTATATTAACCTTTTAAAGAAAATATACCTAAATATAGTAATATTG 842  
 Db 1728 GTCACTACGATCTATTTATGACGAGCTATGTAATATTTGAAAAGTATTTAATTTATTTG 1787  
 Qy 843 TGCATAATCTATATAGACAGACTAATAAATCTTATAGACAGAACCAATATGAATGAG 902  
 Db 1788 CACCAAGACATATACATAAAGATTTGATCCTTTAAAGATCAAGAAAATCATGCGA 1847  
 Qy 903 TATATTTAATGACTATGCAAGATATATGACCATTACTGTATTTAGATACATCTCTCAATT 962  
 Db 1848 TGCTTATTAACATATTCGTCGAAATGACCTTAATTTGATTTGATCTTTCGCAACTTT 1907  
 Qy 963 TTTCTTATATGATATTAATAAGATATAGAGATTCAATTAGAGAAATAGAAAGTAAAGCAT 1022  
 Db 1908 TCTTTTATATGATATATCGTCTTTTCCAGAGAGTAACTAGAAATTAACAGAGAGGT 1967  
 Qy 1023 TAAAG 1027  
 Db 1968 TTATA 1972

## RESULT 13

ADF31301  
 ID ADF31301 standard; DNA; 4896 BP.

AC ADF31301;

XX 12-FEB-2004 (first entry)

DE Bacillus thuringiensis serovar finitimus pf2 DNA clone.

XX Cry26a1; cry28a1; delta-endotoxin; insect pest control;  
 KM transgenic plant; insect resistance; insecticide; gene; ds.

XX Bacillus thuringiensis serovar finitimus.

OS Key Location/Qualifiers

FT CDS 1129..4458  
 /\*tag= a  
 /product= "Cry28a1 delta-endotoxin protein"

XX US2003150018-A1.

XX 07-AUG-2003.

XX 15-JAN-2003; 2003US-00345020.

XX 07-JAN-2000; 2000US-0175158P.

XX 08-JAN-2001; 2001US-00756526.

XX (WOJC/) WOJCIECHOWSKA J A.

XX (LEWIT/) LEWITIN E I.

XX (ZALU/) ZALUNIN I A.

XX (REVIT/) REVINA L P.

XX (CHES/) CHESTUKHINA G G.

XX Wojciechowska JA, Lewitin EI, Zalunin IA, Revina LP,  
 PI Chestukhina GG;

XX WPI; 2003-897623/82.  
 DR P-PSDB; ADF31302.  
 XX New isolated nucleic acid molecule encoding a toxin that is active

PT against insects useful for controlling insect pests or for conferring  
 PT insect resistance in plants.  
 XX  
 PS Claim 7; SEQ ID NO 3; 42pp; English.  
 XX

CC The present invention relates to the isolation of novel cry26a1 and  
 CC cry28a1 delta-endotoxin genes from *Bacillus thuringiensis* spp.  
 CC finitimus. The sequences for the delta-endotoxin polypeptides are also  
 CC disclosed. The invention provides methods for producing the toxins and  
 CC compositions containing the toxins. The methods and sequences of the  
 CC invention are useful for controlling insect pests in transgenic plants to  
 CC confer insect resistance. The present sequence represents a DNA clone  
 CC that contains the coding sequence for cry28a1 delta-endotoxin.  
 XX

XX Sequence 4896 BP; 1727 A; 757 C; 887 G; 1525 T; 0 U; 0 Other;

Query Match 7.4%; Score 163.8; DB 10; Length 4896;  
 Best Local Similarity 55.9%; Pred. No. 3.4e-19;  
 Matches 405; Conservative 0; Mismatches 272; Indels 48; Gaps 3;

Qy 306 AGGTGCTATATATATATCTTTGGTACCTTAATCACTGTCTTTGGCCCGGAGAGACA 365  
 Db 1293 AGGTGTGCTATATATATCAATTTGGAACCTTGGCTCCGCTCTTTGGCTGATCCAGAGGA 1352  
 Qy 366 AGA---CAAAACGATATGACACAAATTTAATAATGGAGAAATTTTGTGTATACACC 422  
 Db 1353 AGATCAAAAATAATTTGGTCACAAATTTATGAAACACGAGAGAACCTTTTAATCAAAAC 1412  
 Qy 423 GTTACAGAAAGCATTAACACGCTAAGTTACAACTTTTGAAGAGTTTGAACAATATT 482  
 Db 1413 AATTTCTACAGCTGTATTAAGAAATAGCATTTAGCTATCTTAATGTGTTTAAGATGATT 1472  
 Qy 483 ACAAGCTATATATACAGCATTTAGATGATGAGAGAAATTAATAAGACTACAGCTCTGG 542  
 Db 1473 AAGCTACTATGAAGAGACATTTATGATGGAAGAGAAATCAAGTCAATATCTGCAG 1532  
 Qy 543 ATTACCAACATCATCAGCATTTACACAAGCTGCTTACTCTTAAATAGATTGAGAA 602  
 Db 1533 ATTGTATCAAGA-----GATTTGAAA 1556  
 Qy 603 TGTTCACAATGATTTTATTTGAGAAATACCTGTTTCAACTTGAATTAATAAGCT 662  
 Db 1557 CGCTCATTTCATTTTGTAGCAATATGCCACAACCTCCACGTTCCAGTTGACACATT 1616  
 Qy 663 ATTACTACTATTTATTCGCAAGCTGCTAATTTTCATTTAATTTATTAACAACAAGTGC 722  
 Db 1617 ATTATTAAGTTGCTATACAGAGCTGCAGAAATTTACATTTGATTTATCATCAAGGTGT 1676  
 Qy 723 TGAATTTGGCTGATGAATGGAATGCAATATACATCTTCCAAATTTGAACCTAATGCTGG 782  
 Db 1677 ACAATTCGCGATCAATGAATGCAATCAACCATTCACAT-----GTTGAA 1727  
 Qy 783 AACATCATGATGACTATATATACTTTTAAAGAAAATATACCTTAATATATAGTAATTTG 842  
 Db 1728 GTCACTACGATCTATTTATGACAGCTATTTGTAATATTTGAAGATATTTAATTTTGG 1787  
 Qy 843 TGCATAATCTATATAGACAGACTTAAATAATCTTGAAGAGAACCAATATGAATGAG 902  
 Db 1788 CACCAAGACATATACATAAAGATTTGAATCACTTAAAGAAATCAAGAAAATCATGCGGA 1847  
 Qy 903 TATATTTAATGACTATGAAAGATATATGACCATTAATCTGTATTTAGATACATCTTCAATT 962  
 Db 1848 TGCTTATTAACATATGCTGCGAGAAATGACCTTAATTTGATTTGATCTTTCGCACTTT 1907  
 Qy 963 TTTCTTATATGATTAATAAGATATATAGATTTCAATGAGAGATTAAGTAAGGAT 1022  
 Db 1908 TCTTTTATGATATATCGTCTTTTCCAGAGAGTGAATCTAGAAATTAACAGAGAGGT 1967  
 Qy 1023 TAAAG 1027  
 Db 1968 TTATA 1972

```
RESULT 14
ADP31306
ID ADF31306 standard; DNA; 4896 BP.
XX
AC ADF31306;
XX
AC ADF31306;
XX
DT 12-FEB-2004 (first entry)
XX
DE Bacillus thuringiensis serovar finitimus pF2 DNA clone.
XX
KM Cry26Aa1; delta-endotoxin; insect pest control;
KM transgenic plant; insect resistance; insecticide; gene; ds.
XX
OS Bacillus thuringiensis serovar finitimus.
XX
FH Key Location/Qualifiers
FT 1129..4458
FT CDS /tag= a
FT /product= "Cry28Aa1 delta-endotoxin protein"
XX
PN US2003154510-A1.
XX
PD 14-AUG-2003.
XX
PF 15-JAN-2003; 2003US-00342821.
XX
PR 07-JAN-2000; 2000US-0175158P.
PR 08-JAN-2001; 2001US-00756526.
XX
PA (WOJC/) WOJCIECHOWSKA J A.
PA (LEMI/) LEMITIN E I.
PA (ZALU/) ZALUNIN I A.
PA (REV1/) REVINA L P.
PA (CHES/) CHESTUKHINA G G.
XX
PI Wojciechowska JA, Lewitin EI, Zalunin IA, Revina LP;
PI Chestukhina GG;
XX
DR MPI: 2003-897757/82.
DR P-PSDB; ADF31307.
XX
PT New delta-endotoxin nucleic acid molecules, cry26Aa1 and cry28Aa1, useful
PT for controlling insect pests and for conferring insect resistance.
XX
PS Claim 7; SEQ ID NO 3; 42pp; English.
XX
CC The present invention relates to the isolation of novel cry26Aa1 and
CC cry28Aa1 delta-endotoxin genes from Bacillus thuringiensis ssp.
CC finitimus. The sequences for the delta-endotoxin polypeptides are also
CC disclosed. The invention provides methods for producing the toxins and
CC compositions containing the toxins. The methods and sequences of the
CC invention are useful for controlling insect pests in transgenic plants to
CC confer insect resistance. The present sequence represents a DNA clone
CC that contains the coding sequence for cry28Aa1 delta-endotoxin.
XX
SQ Sequence 4896 BP; 1727 A; 757 C; 887 G; 1525 T; 0 U; 0 Other;
Query Match 7.4%; Score 163.8; DB 10; Length 4896;
Best Local Similarity 55.9%; Pred. No. 3.4e-19;
Matches 405; Conservative 0; Mismatches 272; Indels 48; Gaps 3;
OY 306 AGGTCCTATTAATATCTTTTGGTACCTTAATCACTGTCTTTTGGCCCGGAGAAACA 365
DB 1293 AGGTGTGTACTTATATATCAATTTGAACTTGGCTCCGCTTCTTGGCTGATCCAGAGA 1352
OY 366 AGA--CAAAACAGATGAGACAAATTTATTTAAATGGAGAAATTTTGTGTATACAC 422
DB 1353 AGATCGAAAAAAATTTGTGTACATTTATGAAACAGGAGAGACCTTTTAAATCAAAAC 1412
OY 423 GTTAAACAGAAAGCATTAACAGCTTAAAGTTACAAACCTTTAGAGATTAGCAAAATAT 482
DB 1413 AATTTCTACAGCTGTAAGAAATGACATTAAGCTCATCTAAATGTTTAAAGATGATTT 1472
```

```
OY 483 ACAAGCTATATATACAGATTTAGATGATGGAGAAAATTAAAAAGATACAGCTCTCG 542
DB 1473 AAGCTACTATGAAGAGATTTATATGATGGAGAGAAATCCAAATGCAAAATATGCGCAG 1532
OY 543 ATTACCAATCATCAGCATTTACAAACAAGCTGCTTGACTTTAAATATGATTTGAGAA 602
DB 1533 ATTGTATCACAGA-----GATTTGAAAA 1556
OY 603 TGTTCACAAATGATTTTATTTGAGAAATATACCTGTTTCCAACTTGAACTTAAAGAGCT 662
DB 1557 CGCTCAATTTCAATTTTGTAAAGCAATATGCCAACTCCAACTCCACGATATGACACATT 1616
OY 663 ATTACTACTATTTATGCGCAAGCTGCTAATTTTCAATTTAATTTATTTACACAGAGTGC 722
DB 1617 ATTATTTAGTTGCTATACAGAAAGCTGCAAAATTTACATTTGATTTATATACAAAGGCT 1676
OY 723 TGAATGGCTGATGAATGGAATGCAATATATCATCTCTACAAATTTGAACCTAATGCTGG 782
DB 1677 ACAATTCGCGGATCAATGGAATGCAATCAACCACTTCACCAAT-----GTTGAA 1727
OY 783 AACATCAGATGACTATTTATTAACCTTTTAAAGAAAATTTACTTAATATAGTAACATTTG 842
DB 1728 GTCAATCAGTACTTATTTATGACGACTATTTGTTATATTTGAAAGATATTTAATTTATTTG 1787
OY 843 TGCAAATACCTATATGAACAGACTAAATAATCTTATAGAGACCAACCAATATGAATGAG 902
DB 1788 CACCAAGACATACCATTAAGATTTGAATCACTTTAAAGAACAGAAAATATCAATGAGGA 1847
OY 903 TATATTTAATGACTATGAGAGATATATGACCATTAAGCTGATTTAGATATCATCTCTCAAT 962
DB 1848 TGCCTATTAACATATGCTGAGAAATGACCTTAATGATTTGATTTGATGCTTGGCAACTTT 1907
OY 963 TTTCTTTATATGATTTAAAAAGATATATAGATTTCAATAGAGAGATATGAAGTAAAGCAT 1022
DB 1908 TCCCTTTTATGATATATGCTGCTTTTCCAAAGAGATGAACTTAAGATTAACAAGAGAGCT 1967
OY 1023 TAAGA 1027
DB 1968 TTATTA 1972
RESULT 15
ABK51132
ID ABK51132 standard; cDNA; 3690 BP.
XX
AC ABK51132;
XX
DT 30-JUN-2002 (first entry)
XX
DE cDNA encoding Bacillus thuringiensis insecticidal protein.
XX
KM Insecticide; transgenic; Coleoptera larvae; ss; gene.
XX
OS Bacillus thuringiensis.
XX
FH Key Location/Qualifiers
FT 187..3690
FT CDS /tag= a
FT /product= "Insecticide protein"
XX
PN JP2002045186-A.
XX
PD 12-FEB-2002.
XX
PF 03-AUG-2000; 2000JP-00236140.
XX
PR 03-AUG-2000; 2000JP-00236140.
XX
PA (SDSB-) SDS BIOTEC CORP.
XX
DR MPI: 2002-356468/39.
DR P-PSDB; AAU80281.
XX
```

PT A protein having insecticidal activity, a DNA encoding said protein, and an agent and a method for preventing harmful organisms.

PS Claim 4; Page 17-18; 19p; Japanese.

XX This invention relates to a crystalline protein comprising a fully defined sequence and the nucleotide sequence encoding this protein. The protein of the invention is an agent for preventing harmful organisms comprising *Bacillus thuringiensis* serovar galleriae SDB502, its mutant or a microbe transformed by a DNA encoding the protein. This microbe can be used to produce a protein containing the protein, or containing a protein having insecticidal activity produced by the SDB502, its mutant or a transformed microbe, a microbe which is transformed by using the above DNA and produces the above protein having insecticidal activity, a plant or a seed transformed by using the above DNA, and *Bacillus thuringiensis* serovar galleriae SDB502 producing a protein comprising and producing a protein showing insecticidal activity. The protein of the invention may have insecticidal activity. The agent is used for preventing Coleoptera larvae. This sequence represents the cDNA encoding the *Bacillus thuringiensis* insecticide protein of the invention

CC Sequence 3690 BP; 1289 A; 599 C; 762 G; 1040 T; 0 U; 0 Other;

CC Query Match 4.5%; Score 98.4; DB 6; Length 3690;

CC Best Local Similarity 43.4%; Pred. No. 7.7e-08;

CC Matches 602; Conservative 0; Mismatches 751; Indels 33; Gaps 2;

```

QY 612 TGAATTTATTCGAAATATCTGGTTTCCAACTTGAATCTTAAACGCTATTACTACC 671
DB 717 TGATTTTGCTCTAAATCCATCTTTTGCATATCTGACAGAGATGACATTTATATC 776
QY 672 TATTATGCGCAAGCTGCTAAATTTTCATTAAATTTATACACAAGTGCTGAATGGC 731
DB 777 AGTATATGACACACAGAGATTTTACATTTCTATTATACAGATGCTTCATTTTGG 836
QY 732 TGATGANTGAGATGCAATATACATCTTCAAAATGAACTTAAGTGCGTAACATGCA 791
DB 837 AGCAGAGTGG-----GGATTTCACACAGAGAAATTTTC 869
QY 792 TGACTATTATTAATCTTTTAAAGAAATATATCTAAATATAGTAATTTTGCAGAAATAC 851
DB 870 CACATTTTATATGTCGACGAGACGACCGCCCAATCTCGGATTTATGTGTTAAAGTG 929
QY 852 CTATAGAACAGAGCTAAATAATCTTAGAGCAACCAATATGAAATGAGATATATTTAA 911
DB 930 GTATATACACTGCTTAATTAATTAAGGTACGAATGCTCAAGTTGGCTGAAGTATCA 989
QY 912 TGACTATGAGATATATGACCATTAATCTATGATATACATCTCTCAATTTTCTTATA 971
DB 990 CCAATTCGGAAGAGAAATGACATTATGTAATTTAGTACGTTATTTCCAAACTA 1049
QY 972 TGATATATAAAGATATAGATTCATATAGAGAAATAGAAATTAAGCATTAAGAAATGA 1031
DB 1050 TGACACACGTAACGTAATCCATCGAAACACAGCCCACTTACAGGAAAGTATACGA 1109
QY 1032 ACTCACAGAGAAATTTATACACTGAAATTAATTTTGATGCTTCTCAACTTGAAGT 1091
DB 1110 TCCAAATGATATTTAACAAGAAACAAGTGATTTTGTAGCGTGTGTCACTTAAACAG 1169
QY 1092 TCAACCCAACTACTAGATGAGATATTAATTAACAGTGCAGTTTAAATTTATTTTC 1151
DB 1170 TGATATTTCTTTTTCAGAAAGTGAAGCGCTGTAATTCGTTACACACACTTATTTGAT 1229
QY 1152 ATTTTGAACAATTTATTTTATACAGAAATACAAATTTCCGGAATCGTTAGTTGG 1211
DB 1230 ACTCAGTAATTAATTAATTTTATACACAAGCGGGGCTTCCCTTGAAATATACGAATA 1289
QY 1212 TATTTTAAATGATGACACTACT-----TATAGCAATATCTATTAACGAACTTTATA 1265
DB 1290 CCTGAATATTTGGTGAAGACATTTCTATTAATTAATTAATTAATGATGCTCATCAGCAT 1349
QY 1266 TGGAGAAAGAACAGTTTCAACCAACAAATAAAGACATTTGAATCTTATTAAGT 1325

```

```

DB 1350 AGAAGTAATTTACGGTACGATTTACTTTAACAATAATCAAGATATTTAGCAATTA 1409
QY 1326 TTCATATGTAAGTATGACAAATCACTCTGTTTCCCTATTTCACACACTTTATAT 1385
DB 1410 GGAATATCTTTCAGGTGATCATTTAGGGGCGGATTTAGTAATTAATTAACGACAGTATA 1469
QY 1386 TAATCAATTTGAACCTTATTTTAAATGGCTCATCTAACACACACTCAAAATTTACAGAG 1445
DB 1470 TGGAGTTCCTGACGTAAGTTTACACTCTTGACCAAGAAATACAGGATCAGATTCAGTTGG 1529
QY 1446 AGGGCTTTATCTAATATATCAAAACACAACTTTTTCATTTTCTGAAAAAAGACTG 1505
DB 1530 AGGTTTACGTACTCAAAACACATACATTAAGCAAGTATGTACAAAATTTCAATAC 1589
QY 1506 CAATCTAGTATTTGATTCAGAGTTGTTTCCAAACTTTAATTAATAGTATATTTATC 1565
DB 1590 GATTAAGAAATCCCTCAGAGATGAGCCACTAGTAGAGGGATATGCCATGATATATC 1649
QY 1566 CCAATTTTCAATTTTATTTACTTATTTCTATGATTTGATTCAGCTACAAATTTAGTAC 1625
DB 1650 TCAATATCACTCTTAATTTCTTTTCTAAGAAATGCTAGTAGTCTGCTAGATATGGCAATCT 1709
QY 1626 AGGTGATTTAGATGACACAGACAGTAGTGTGATATATATATGCAATATCAGATAAAT 1685
DB 1710 CCTGTATTTGCTTGGACACATGGAATGCGGATGTTCAAAATACAGTTTATTCAGATTA 1769
QY 1686 AATTACATATATCCACCAATCAAAAGGTAACTTTGATATCAAACTTAAAGTAAATGA 1745
DB 1770 AATTACTCAATATCAGATGTGTAAGGACATATTTAGTTTCAAGTACTACTGTTATTA 1829
QY 1746 AGAAGCTGTGATACAGAGAGAAACTTGTTTATTTTCAAAAGTCAAGGGGCTTTAGAAAT 1805
DB 1830 AGGTCCTGGAATTAACAGAGAGCAATATCTTAAAGAAACAAAGTAGTGTCGTTAAGCTTA 1889
QY 1806 TACATGGAACCTCTTAATCTACAAATCTTATTTCAATAGACTTGATATGCTACAAA 1865
DB 1890 TACTAGTGTCTGTAAATATCACCATTATACAAAGATATCGTCAAGAAATAGCTTATGC 1949
QY 1866 TGGTGTGAAATATCTTCTCTATATATCTCTTACAAATACAGAGTAAATAGGAATACC 1925
DB 1950 TTTCTACTATCACTTACAGCTTTTGTGTAACAATTTTCGAATCGCATTTTACTTATTA 2009
QY 1926 ACCTCAAGACTCAACAACCTTTTCTGTGTAACAATTTATTAATTTTACATACGAGGA 1985
DB 2010 TGTTAATTAACATGATTAAGGAGATGATTTTAACATTTAATATACATTTGATTAAC 2069
QY 1986 TTTTGG 1991
DB 2070 TATTTGG 2075

```

Search completed: December 11, 2005, 14:34:14  
Job time : 1341 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 11, 2005, 13:36:36 / Search time 1134 Seconds  
(without alignments)  
11272.709 Million cell updates/sec

Title: US-10-783-417-1  
Perfect score: 2208  
Sequence: 1 atgacacaaataacgataa.....atcacacaaacacttga 2208

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapect 1.0

Searched: 5883141 seqs, 28421725653 residues

Total number of hits satisfying chosen parameters: 11766282

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : GenEmbl:\*  
1: gb\_ba:\*  
2: gb\_in:\*  
3: gb\_env:\*  
4: gb\_com:\*  
5: gb\_ov:\*  
6: gb\_pat:\*  
7: gb\_ph:\*  
8: gb\_pr:\*  
9: gb\_ro:\*  
10: gb\_strs:\*  
11: gb\_sy:\*  
12: gb\_un:\*  
13: gb\_vl:\*  
14: gb\_hcg:\*  
15: gb\_pl:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2208	100.0	2208	6	CQ868305 Sequence
2	1903.4	86.2	2235	6	CQ868307 Sequence
3	1794.4	81.3	2085	6	CQ868309 Sequence
4	346.6	15.7	4451	6	I08083 Sequence 1
5	346.6	15.7	4451	6	I09103 Sequence 1
6	345.4	15.6	3543	1	BTTOXD1
7	345	15.6	4253	1	BACISRH4
8	345	15.6	4934	6	E01676
9	345	15.6	127923	1	BTBTOXIS
10	222.6	10.1	6009	1	BTBTOXIS
11	189.6	8.6	127923	1	BTBTOXIS
12	188	8.5	3753	1	BACMSQH
13	184.2	8.3	434	1	DO078744
14	178.2	8.1	3756	6	E00614
15	171.6	7.8	1993	6	I03578
16	166.8	7.6	4959	1	AF285775
17	163.8	7.4	4896	1	AF132928
18	157	7.1	3752	1	AB125059

19	143.8	6.5	3642	1	AB193814
20	133.6	6.1	3644	1	BTB251978
21	98.4	4.5	3435	1	AB089299
22	98.4	4.5	3504	6	BD133574
23	98.4	4.5	3690	6	BD133575
24	98	4.4	1738	6	I03580
25	97	4.4	104992	14	AC005504
26	97	4.4	169546	14	AC004157
27	97	4.4	250421	2	AE014849
28	97	4.4	349980	6	AX344555
29	94.6	4.3	15548	6	AX347057
30	94.6	4.3	133877	14	AC120883
31	92.4	4.2	86826	2	PFMAL3P5
32	91.8	4.2	250029	2	AE014839
33	91.8	4.2	313050	2	PPA929352
34	90.8	4.1	3668	1	DB8381
35	90.4	4.1	8056	6	AX599046
36	86.2	3.9	254050	2	PPA929358
37	86	3.9	4660	6	CS083843
38	86	3.9	90550	8	AL592166
39	86	3.9	176898	8	AC117569
40	85.8	3.9	3444	6	AR050192
41	85.8	3.9	3444	6	AR050194
42	85.8	3.9	3444	6	AR060320
43	85.8	3.9	3444	6	AR060322
44	85.8	3.9	3444	6	I19463
45	85.8	3.9	3444	6	I19465

ALIGNMENTS

RESULT 1  
LOCUS CQ868305 2208 bp DNA linear PAT 13-SEP-2004  
DEFINITION Sequence 6 from Patent WO2004074462.  
ACCESSION CQ868305  
VERSION CQ868305.1 GI:51998351  
KEYWORDS  
SOURCE Bacillus thuringiensis  
ORGANISM Bacillus thuringiensis  
Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus; Bacillus cereus group.  
REFERENCE 1  
AUTHORS Carozzi, N., Hargiss, T., Koziel, M.G., Duck, N.B. and Carr, B.  
TITLE Delta-endotoxin genes and methods for their use  
JOURNAL Patent: WO 2004074462-A 6 02-SEP-2004;  
Athenix Corporation (US)  
FEATURES  
Source  
1..2208  
/organism="Bacillus thuringiensis"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:1428"  
1..2208  
/note="unnamed protein product"  
/codon\_start=1  
/transl\_table=11  
/protein\_id="CAH33948.1"  
/db\_xref="GI:51998352"  
/translation="MNQNDNNEYEIIIDSHSPYFPNNNSDRPYTNNPQLONT  
NYKEMLNMCQNTGDNFEPASADITFAASAGTIVSGTLAGIGLTSISGPGLI  
GALISFGLITFMPAGEODKVTWGTGKGLIFVDPLESIKQLTLEGROJ  
LOSNTLDDMRKIKRLQAPGLPPSALQQAALTKIRFENVHNDIFIEIGFOLEY  
KTLILPIYAQANHLNLTQCGALADBNMADLHPSQIEPNAGTSDDVYKIKENIPK  
YSNVCANTYRTGLKLRDEBPMKKSIRNDYRRTITVLTLDITSQPSLDIKRYRDSIG  
GIEVKGINKELTREIYVTEINFDRLPQLRVGNPLATWENVTLRASFLEDEPIFY  
TEENTNGRLVGINRDAPYNSNTITETLYGERTGSPPTKTIREFESKYVAIVDRS  
PVPSPIDPFIINOIEIYNGSSNNTLYSGGSLSNYQNTTIPQFPKQCNVLDIP  
GSPNPNVSHIILSHPSLFTYSYVIGLOLIDLVGVGMTSSDRAVAIDKTIIMI  
PAIKGNLDNPKSVKIRGEGHGVNLYLOSGRLTEICETNSQSVIRLRVYANGA  
GNTLPNLSITIPGVIGIPPORLANTPSTRTNNQIQGDGPGYFGPSPYTLPLNRIPF  
IFNRADVNSIILIDKLEFIPTISSMQRNKKQLETTIQTKINFPFNHTKYL"

ORIGIN

Query Match 100.0%; Score 2208; DB 6; Length 2208;  
Best Local Similarity 100.0%; Pred. No. 4,96-250;  
Matches 2208; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGAATCAAAATTAAGCATATATACGAATATGAAATTTATGATTCGCATACCTCACTTAT 60  
DB 1 ATGAATCAAAATTAAGCATATATACGAATATGAAATTTATGATTCGCATACCTCACTTAT 60

QY 61 TTTCCGAAACAGAAACAGTATGATTTCTAGATACCTTTACACAAATTAATCCAAATCAACA 120  
DB 61 TTTCCGAAACAGAAACAGTATGATTTCTAGATACCTTTACACAAATTAATCCAAATCAACA 120

QY 121 TTACAAAACAAATTAACAAAGAGTGGCTCAATATGTCTCAAGGGAATTAACAATATGCT 180  
DB 121 TTACAAAACAAATTAACAAAGAGTGGCTCAATATGTCTCAAGGGAATTAACAATATGCT 180

QY 181 GATAATTTTCAGAGACATTTGCTAGTCTGATACAAATTTGCTGACAGTATGAGCATATTT 240  
DB 181 GATAATTTTCAGAGACATTTGCTAGTCTGATACAAATTTGCTGACAGTATGAGCATATTT 240

QY 241 GTATCCGGTACTCTGTATAGCCGGTATAGTGGGCTCACTTCTATATCCGACCGATAGA 300  
DB 241 GTATCCGGTACTCTGTATAGCCGGTATAGTGGGCTCACTTCTATATCCGACCGATAGA 300

QY 301 ATATATAGTGTCTATATATATATCTTTTGGTACCTTAATCACTGTCTTTTGGCCCGGGA 360  
DB 301 ATATATAGTGTCTATATATATCTTTTGGTACCTTAATCACTGTCTTTTGGCCCGGGA 360

QY 361 GAACAGACAAACAGTATGAGCACAATTTTAAATGGGAGAAATTTTGTGATATCA 420  
DB 361 GAACAGACAAACAGTATGAGCACAATTTTAAATGGGAGAAATTTTGTGATATCA 420

QY 421 CCGTTAACAGAAAGCATATAACAGCTAAAGTTACAACTTTTGAAGAGATTTTGAACAATA 480  
DB 421 CCGTTAACAGAAAGCATATAACAGCTAAAGTTACAACTTTTGAAGAGATTTTGAACAATA 480

QY 481 TTACAAAGCTATATATACAGCATTTAGATGATGGAGAAATTTTAAAGAGCTACAGCTCTT 540  
DB 481 TTACAAAGCTATATATACAGCATTTAGATGATGGAGAAATTTTAAAGAGCTACAGCTCTT 540

QY 541 GGATTACCAACATCATCGACATTTACAACAGCTGCTTGACTCTTAAATTCGATTTAG 600  
DB 541 GGATTACCAACATCATCGACATTTACAACAGCTGCTTGACTCTTAAATTCGATTTAG 600

QY 601 AATGTTCAACATGATTTTATTCGAGAAATACCTGGTTTCCAACCTTGAATTAATAACG 660  
DB 601 AATGTTCAACATGATTTTATTCGAGAAATACCTGGTTTCCAACCTTGAATTAATAACG 660

QY 721 GCTGAATTTGGCTGATGAAATGGAATGAGATATACATCCCTTCAACAAATTTGAACCTTAATGCT 780  
DB 721 GCTGAATTTGGCTGATGAAATGGAATGAGATATACATCCCTTCAACAAATTTGAACCTTAATGCT 780

QY 781 GGAACATCAGATGATCTATTTATTAACCTTTTAAAGAAATATATCTTAAATATATAGTAACAT 840  
DB 781 GGAACATCAGATGATCTATTTATTAACCTTTTAAAGAAATATATCTTAAATATATAGTAACAT 840

QY 841 TGTGCAATTAACCTATATGAAACAGGACTTAAATTTCTTGAAGAGAAACCAATATGAAATGG 900  
DB 841 TGTGCAATTAACCTATATGAAACAGGACTTAAATTTCTTGAAGAGAAACCAATATGAAATGG 900

QY 901 AGTATATTTAATGATCTATGAAAGATATATGACATTTAGTATGATATCAATCTCTTAA 960  
DB 901 AGTATATTTAATGATCTATGAAAGATATATGACATTTAGTATGATATCAATCTCTTAA 960

QY 961 TTTTCTTTATATGAT 1020  
DB 961 TTTTCTTTATATGAT 1020

QY 1021 ATTAAGAATGAACTCACAGAGAAATTTATACAACTGAAATTAATTTTGAATCGCTTCT 1080  
DB 1021 ATTAAGAATGAACTCACAGAGAAATTTATACAACTGAAATTAATTTTGAATCGCTTCT 1080

QY 1081 CAACCTTGAAGTCAACCCAACTCTAGATGAGATATATATTTTAAACAGGTGCAAGTTT 1140  
DB 1081 CAACCTTGAAGTCAACCCAACTCTAGATGAGATATATATTTTAAACAGGTGCAAGTTT 1140

QY 1141 AAATTAATTTTCAATTTTGAACAAATTTATTTTATACAGAAATTAACAAATTTGCGAAT 1200  
DB 1141 AAATTAATTTTCAATTTTGAACAAATTTATTTTATACAGAAATTAACAAATTTGCGAAT 1200

QY 1201 CGTTAGTGGTATTTCTAATTCGATGACCTTACTTATGCAATACTATTAATCGAAT 1260  
DB 1201 CGTTAGTGGTATTTCTAATTCGATGACCTTACTTATGCAATACTATTAATCGAAT 1260

QY 1261 TTATATGAGAAAGAAACAGGTTCAACCAACAAACAAATTAAGACCAATTTGAATCTTAT 1320  
DB 1261 TTATATGAGAAAGAAACAGGTTCAACCAACAAACAAATTAAGACCAATTTGAATCTTAT 1320

QY 1321 AAAGTTTCAATTTGATCTGATATGACAAATCACTCTGTGTTCCCTATTTCAACACACTTT 1380  
DB 1321 AAAGTTTCAATTTGATCTGATATGACAAATCACTCTGTGTTCCCTATTTCAACACACTTT 1380

QY 1381 ATATATTAATCAAAATGGAACCTTTATTTAAATGAGTCACTTAACAACACATCAATATTTCA 1440  
DB 1381 ATATATTAATCAAAATGGAACCTTTATTTAAATGAGTCACTTAACAACACATCAATATTTCA 1440

QY 1441 GCAGAGAGGCTTTTATCTAATTTATCAAAACCAACTTTTTCATTTCTCTAGAAAAAA 1500  
DB 1441 GCAGAGAGGCTTTTATCTAATTTATCAAAACCAACTTTTTCATTTCTCTAGAAAAAA 1500

QY 1501 GACTGCAATCTAGTATATGATTCAGAGTGTTCACGAAACTTTTATATATCTATGATAT 1560  
DB 1501 GACTGCAATCTAGTATATGATTCAGAGTGTTCACGAAACTTTTATATATCTATGATAT 1560

QY 1561 TTATCCCATTTTTCATTTATTTACTATTTCTTATGATATGATATGATATGATATGATATGATAT 1620  
DB 1561 TTATCCCATTTTTCATTTATTTACTATTTCTTATGATATGATATGATATGATATGATATGATAT 1620

QY 1621 GATACAGGTGATATGAGATGGAACACAGTATGATATGATATGATATGATATGATATGATAT 1680  
DB 1621 GATACAGGTGATATGAGATGGAACACAGTATGATATGATATGATATGATATGATATGATAT 1680

QY 1681 AAAATTAATTAATGATATCCAGCAATCAAAAGTAAACAATCTTGATACAACTCTTAAGGTA 1740  
DB 1681 AAAATTAATTAATGATATCCAGCAATCAAAAGTAAACAATCTTGATACAACTCTTAAGGTA 1740

QY 1741 ATGGAAGGACCTGTCATACAGAGGAAACCTGGTTTATTTTCAAAAGTCAAGGCGTTTA 1800  
DB 1741 ATGGAAGGACCTGTCATACAGAGGAAACCTGGTTTATTTTCAAAAGTCAAGGCGTTTA 1800

QY 1801 GAAATTAATCATGTAATCTCTTAATTTTACAAATCTTATTTTCAATTTTGAATGCT 1860  
DB 1801 GAAATTAATCATGTAATCTCTTAATTTTACAAATCTTATTTTCAATTTTGAATGCT 1860

QY 1861 ACAATATGCTGGAATATCTCTTCTAATATATCTCTTCAATATCAAGAGATATAGTA 1920  
DB 1861 ACAATATGCTGGAATATCTCTTCTAATATATCTCTTCAATATCAAGAGATATAGTA 1920

QY 1921 ATACACACCTCAACAGCTCAACACACTTTTCTGATCAAAATTTATATATTTTCAATATC 1980  
DB 1921 ATACACACCTCAACAGCTCAACACACTTTTCTGATCAAAATTTATATATTTTCAATATC 1980

QY 1981 GGAGATTTTGGGTATTTTCAATTTTCAAGTATAGTATACATTAACCTTTAAATTCGAACATA 2040  
DB 1981 GGAGATTTTGGGTATTTTCAATTTTCAAGTATAGTATACATTAACCTTTAAATTCGAACATA 2040

QY 2041 CCATTTATATTTATATCGTGAGATATCAATTTCAATTTTATCATTTGATATTAATTTGAA 2100  
DB 2041 CCATTTATATTTATATCGTGAGATATCAATTTCAATTTTATCATTTGATATTAATTTGAA 2100

QY 2101 TTTATACCAATTAATCTTCTCTATGACCAAAATTAAGAAAAACAAATTTAGAAATCTATC 2160

DB	2101	TTTATGCAATTTACTTCTCTATGACCAAAATAGAGAAAAACAAAATTGCAACTATC	2160
Qy	2161	CAACCAAAATAATACATTTTTCACAATGATCAAAAACACTTTGA	2208
Db	2161	CAACCAAAATAATACATTTTTCACAATCATCAAAAACACTTTGA	2208
RESULT 2			
Q0868307			
LOCUS	C0868307	2235 bp	DNA
DEFINITION	Sequence 8 from Patent WO2004074462.		linear
ACCESSION	C0868307		
VERSION	C0868307.1	GI:51998353	
KEYWORDS			
SOURCE			
ORGANISM	Bacillus thuringiensis		
	Bacillus thuringiensis		
	Bacteria: Firmicutes; Bacillales; Bacillaceae; Bacillus; Bacillus		
	cereus group.		
REFERENCE			
AUTHORS	Carozzi, N., Hargiss, T., Koziel, M. G., Duck, N. B. and Carr, B.		
TITLE	Delta-endotoxin genes and methods for their use		
JOURNAL	Patent: WO 2004074462-A 8 02-SEP-2004;		
FEATURES	Athenix Corporation (US)		
source	Location/Qualifiers		
	1..2235		
	/organism="Bacillus thuringiensis"		
	/mol_type="unassigned DNA"		
	/db_xref="taxon:1428"		
	1..2235		
	/note="unamed protein product"		
	/codon_start=1		
	/transl_table=1		
	/protein_id="CAH33949.1"		
	/db_xref="GI:51998354"		
	/translation="MGNNNNEVEIILDSKLSIYPSNRNIDHSRIPTYNNPNQLQNTN		
	YKELANLQCGVQYGDNFEFTASADTIAVASAGTVSGTLLAGIGLTSISGPIIG		
	AIIISFGTLLTFWMPAGBODKTWTFQIKMEIIVDPDLPIESIKQLKDLREGRLIK		
	QSVATLDDMRKLRLOAPGLPPSSALQOALTLKRPENVANDPIREIPGQLETK		
	TLILPYAQAANFHLNLQOAGLADENDMLHPSOIEPNAGTSDDVYKLLKENIPEY		
	SNVCAVYREGNLRLNRPNRMSIENDRYRMYTITVLDPTAOPSFYDIKRYDSIER		
	IGGKTELREIYITEINFDRLLTYLLEIOPNLALMEVNLTSGLRISFLBELIPTYK		
	ETVGNRLVGLANRRSTYATGTETIYGERGTPPTYTLIPESYKVSIVTDRLVPTP		
	SPFNPIYFNIQLELYLNNSPSKLTYSAGGSLNDKKTDFQPPVKDDCKIINPC		
	LPSNYSKILISQPSLFNYSYKIGLALNIYTGALGWTHSSVNNVNAISDKITIMI		
	IKGSLDPMNSKVIKRGPHGTGKNIYVLIOSQRLITCGKPNSTGSYVRLRYATNGA		
	TLPIISLITPESVIGIIPQRLNNTSGRNRYNLQYGDGYQYQPESTVILPLNRNIP		
	PLTF		
	NRAVDSILIIDKIEFIPIITSSVNRNREKQLETIQTKINTFTNHTKNTLNTA		
	YDID"		
ORIGIN			
Query Match	86.2%; Score 1903.8; DB 6; Length 2235;		
Best Local Similarity	92.2%; Pred. No. 8e-249; Indels 9; Gaps 2;		
Matches 2030; Conservative	162; Mismatches		
Qy	11	ATAACGATTAATAAGCAATATGAATTAATGATTCGCATACCTCACTTATTTCCGAACA	70
Db	8	AAAAATATAATAAGCAATATGATATGATTCGATTCGAAGATTTATCTTATCCTCTAACA	67
Qy	71	GAACAGTAATGATTTAGATACCCCTTACACAAATAATCCAAATCAACCATTCACAAAACA	130
Db	68	GAATATGATCACTTATGATACCCCTTACACAAATAATCCAAATCAACCATTCACAAAACA	127
Qy	131	CAAAATTCACAAAGAGTGGCTCAATATGTGTCAAGGCAATACCAATATGTGTATTAATTCG	190
Db	128	CAAAATTCACAAAGAGTGGCTCAATATGTGTCAAGGCAATACCAATATGTGTATTAATTCG	187
Qy	191	AGACATTTGCTAGTGCCTGATCAATTTGGTGGAGTTAAGCAGATCAATTTGATCCGGTA	250
Db	188	AGACATTTGCTAGTGCCTGATCAATTTGGTGGAGTTAAGTGCAGATCAATTTGATCCGGTA	247
Qy	251	CTCTGTTAGCCGCTATAGTGGGCTCACTTCTATATCCGACCAATGGAATAATAGGTG	310

D	248	CTCTGTTAGCCGGATATAGTGGGCTCACTTCTATATCCGACCGATAGAAATTAATAGGTG	307
Q	311	CTATATATATATCTTTTGGTACCCCTAATCACTGTCTTTTGGCCCGGGAGAACAGACA	370
D	308	CTATATATATATCTTTTGGTACCCCTAATCACTGTCTTTTGGCCCGGGAGAACAGACA	367
Q	371	AAACAGTATGGACCAATTTATTTAAATGGAGAAATTTTGTGTGATCACCCGTTAACG	430
D	368	AAACAGTATGGACCAATTTATTTAAATGGAGAAATTTTGTGTGATCACCCGTTAACG	427
Q	431	AAAGCATTAACACACTAAAGTTACAAACTTTAGAGAGATTAGACAAATATTAACAAGCT	490
D	428	AAAGCATTAACACACTAAAGTTACAAACTTTAGAGAGATTAGACAAATATTAACAAGCT	487
Q	491	ATAATACAGCATTTAGATGATTGGAGAAATTTAAAGAATAACAAGCTCTGGATTACAC	550
D	488	ATAATACAGCATTTAGATGATTGGAGAAATTTAAAGAATAACAAGCTCTGGATTACAC	547
Q	551	CATCATCAGCATTTACACACAGCTGCCCTTGACTCTTAATAATCGATTGGAAATGTTCACA	610
D	548	CATCATCAGCATTTACACACAGCTGCCCTTGACTCTTAATAATCGATTGGAAATGTTCACA	607
Q	611	ATGATTTTAAATTCGAGAAATACCTGGTTCCAACTGTGAACCTTATTAACCGCTATTAAC	670
D	608	ATGATTTTAAATTCGAGAAATACCTGGTTCCAACTGTGAACCTTATTAACCGCTATTAAC	667
Q	671	CTATTATATGGCAGCAAGCTGTCTAATTTTCATTTAAATTTATTAACAAGAGTCTGAATTTG	730
D	668	CTATTATATGGCAGCAAGCTGTCTAATTTTCATTTAAATTTATTAACAAGAGTCTGAATTTG	727
Q	731	CTGATGATGGAATGCAAGATATACATCTTTCACAAATTTGAACCTAATGCTGGAAATCAG	790
D	728	CTGATGATGGAATGCAAGATATACATCTTTCACAAATTTGAACCTAATGCTGGAAATCAG	787
Q	791	ATGACTATTTAACTTTTAAAGAAATATPACTAATATATAGTATCTATGTGCCAATA	850
D	788	ATGACTATTTAACTTTTAAAGAAATATPACTAATATATAGTATCTATGTGCCAATA	847
Q	851	CCTATTAAGAACGCACTAAAAAATCTTGAAGCAGAACCAATATGAAATGGAGTATATTTA	910
D	848	CCTATTAAGAACGCACTAAAAAATCTTGAAGCAGAACCAATATGAAATGGAGTATATTTA	907
Q	911	ATGACTATCGAAGATATATGACCAATTACTGTATTAGATACCATCTCTCAATTTCTTTAT	970
D	908	ATGACTATCGAAGATATATGACCAATTACTGTATTAGATACCATCTCTCAATTTCTTTAT	967
Q	971	ATGATATTAAGAAAGTATAGAGATTCATTAAGAGAAATAGAAAGTAAAGCATTTAAGAAATG	1030
D	968	ATGATATTAAGAAAGTATAGAGATTCATTAAGAGAAATAGAAAGTAAAGCATTTAAGAAATG	1021
Q	1031	AACCTACAGAGAAATTTATATCAACTGAATAATTTTGTGATCGTCTCTCAACTTTAGAG	1090
D	1022	AACCTACAGAGAAATTTATATCAACTGAATAATTTTGTGATCGTCTCTCAACTTTAGAG	1081
Q	1091	TTCAACCCAAATCTAGCTACGATGGAATATATTTAAACAGCTGCAGAGTTTAAATTAATTT	1150
D	1082	TTCAACCCAAATCTAGCTACGATGGAATATATTTAAACAGCTGCAGAGTTTAAATTAATTT	1141
Q	1151	CATTTTATGAACAATTTATTTTATATACAGAAATACAAATTTTGGGAATCGTTTAGTG	1210
D	1142	CATTTTATGAACAATTTATTTTATATCAAAAAATGAACGTACGGGAATCGTTTAGTG	1201
Q	1211	GTATTTTCAATTCGATGATGACCTACTATATAGCAATACATTAACGTAACTTATATAGGAG	1270
D	1202	GTATTTGGAATTCGATATAGATCTACTATATAGCTACGACAGGAATCTGAATATATATAGGAG	1261
Q	1271	AAAGAACAGGTTCAACCCACAAACAAATAAGAACATTTGGAATCTTATTAAGCTTTCAA	1330
D	1262	AAAGAACAGGTTCAACCCACAAACAACTTAATATACATTTGGAATCTTATTAAGCTTTCAA	1321
Q	1331	TTGTATCTGATAGCAATCACTCTCTGTCTTCCCTATTTC---AACCACTTTATTAATTA	1387
D	1322	TTGTATCTGATAGCAAGTATCACTCTCTTCCCTTTCTTAACATATTAATTAATTAATTA	1381

QY 1388 ATCAATTTGAATTTTAAATGGCTCATTAACAACACTCAAAATATTTCAGCAGAG 1447  
DB 1382 ATCAAAATTTGAATTTTAAATTAATTCACCTAGTAATAAATTAACATATTTCAGCTGGGG 1441  
QY 1448 GGTCTTATTCATTAATTAACAAAACAATTTTTCATTTTCCTAGGAAAAAAGACTGCA 1507  
DB 1442 GGAATTTATTCATTAATTAACAAAACAATTTTCATTTTCCTAGGAAAAAAGACTGTA 1501  
QY 1508 ATCTAGTTATTCATTCAGAGTGTTCACCAAACTTTAATACATATGCTATTTTATCC 1567  
DB 1502 AACCAATTTATTCATTCAGAAATTTTATTCACAGCTAATAATAGTATATGCTATTTATCC 1561  
QY 1568 ATTTTTCATTTATTTATCTTATTCCTATGATGATTAAGCTTACCAATATTAGATACAG 1627  
DB 1562 AGTTTCTTTATTTATTAATTTCTAATAAAATTTGATTTAGCCCTAATATATTAATATCAG 1621  
QY 1628 GTGTATTAAGATGACACACAGTAGTGTGTATGATATTAATGCAATATACATTAATAA 1687  
DB 1622 GTGCTATTAAGATGACACACAGTAGTGTATTAATGCAATATACATTAATAA 1681  
QY 1688 TTACAATGATCCACGACATCAAAAGTAAACAATCTTGATACAACTTAAGGTAATTGAG 1747  
DB 1682 TTACAATGATCCACGACATCAAAAGTAAACAATCTTGATACAACTTAAGGTAATTGAG 1741  
QY 1748 GACCTGTGATACAGAGGAAACTTGTTTATTTTCAAAAGTCAAGGCTTTAGAAATTA 1807  
DB 1742 GACCTGTGATACAGAGGAAACTTGTTTATTTTCAAAAGTCAAGGCTTTAGAAATTA 1801  
QY 1808 CATGTGAAATCCCTTAATCTACACAACTTTATTTTCAATTAAGCTTGCATGCTCAATG 1867  
DB 1802 CATGTGAAATCCCTTAATCTACACAACTTTATTTTCAATTAAGCTTGCATGCTCAATG 1861  
QY 1868 GTGCTGAAATACTCTTCTAATATATCTCTTAACAATACAGAGTATAGAAATACAC 1927  
DB 1862 GTGCTGAAATACTCTTCTAATATATCTCTTAACAATACAGAGTATAGAAATACAC 1921  
QY 1928 CTCAACGACTCAACAACACTTTTCTGTGTAACAATTTATTAATTTTCAATACGAGATT 1987  
DB 1922 CTCAACGACTCAACAACACTTTTCTGTGTAACAATTTATTAATTTTCAATACGAGATT 1981  
QY 1988 TTGGGATTTTCCAAATTTCCAAATACAGTAACAATTAACCTTTAATGCAAACTTACATTTA 2047  
DB 1982 TTGGGATTTTCCAAATTTCCAAATACAGTAACAATTAACCTTTAATGCAAACTTACATTTA 2041  
QY 2048 TATTTAATCGTACAGATGATCAATCAATTTTATCAATGATTAATTAATGATTTATAC 2107  
DB 2042 TATTTAATCGTACAGATGATCAATCAATTTTATCAATGATTAATTAATGATTTATAC 2101  
QY 2108 CAATTTACTTCTCTATGACCAAAATAGAGAAAAAATAAATTGAAAGTATCCAAACA 2167  
DB 2102 CAATTTACTTCTCTATGACCAAAATAGAGAAAAAATAAATTGAAAGTATCCAAACA 2161  
QY 2168 AATTAATACATTTTTCACAAATCATCAAAACAACCTTTGA 2208  
DB 2162 AATTAATACATTTTTCACAAATCATCAAAACAACCTTTGA 2202

RESULT 3  
C0868309 2085 bp DNA linear PAT 13-SEP-2004  
LOCUS C0868309  
DEFINITION Sequence 10 from Patent WO2004074462.  
ACCESSION C0868309  
VERSION C0868309.1 GI:5198355  
KEYWORDS  
SOURCE  
ORGANISM  
Bacillus thuringiensis  
Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus; Bacillus  
cereus group.  
REFERENCE  
1 Carozzi, N., Hargiss, T., Kozief, M. G., Duck, N. B. and Carr, B.  
AUTHORS Delta-endotoxin genes and methods for their use  
TITLE Patent: WO 2004074462-A 10 02-SEP-2004;  
JOURNAL

FEATURES  
source Athenix Corporation (US)  
Location/Qualifiers  
1..2085  
/organism="Bacillus thuringiensis"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:1428"  
1..2085  
/note="unnamed protein product"  
/codon\_start=1  
/transl\_table=1  
/protein\_id="CAH33950.1"  
/db\_xref="GI:5198355"  
/translation="MCGATQYQDNEPTASADTTAAVSAAGTIVSGTLAIGGTSI  
SGPIGLIGALISFGTLIVFWPAGSDKTVMQPIKMSI FVDPLTBSIKQLKLT  
LBSGRQILQSYNTALDDMKLKRLOAQPSSALQOALTLKIREPNVNDRIEIP  
GFQLETKTLLEIYAOAAFNHLLOQGAELDENADIDHPSQIBENAGTSDYVL  
LKENIPKYSNYCANTREGANKLRBNPNMSI.FNDYRMYTITVDTIOFSPYDIK  
RYDSIGRIGGIGITELTREIYTEINDETLIEIOPNLAIMENLITRSLRIPSPFD  
ELI.FYTKNEYGRVLGIANRNSSTATTTEIITGERTPPPTTKLIPRESKYSLV  
TDRQVPTSPFPNIYFTINQILBYLNNSPKNLTYSAGNLSDKTTDQFPVKDC  
KPIINPCLPSVNSYSHILSQFSLFNVSYKIGALNLLYGAIGMTHISVNRNALSD  
KIITMT.PAIKGNLIDTNSKVI.BSGPHQGNLVYLOSGRLEICRTENSTQSYIIRL  
YATNGAGNTLNPISLTIPGIVIGIPORLNNTPSGTNNNNI.OYGPFGFQPSYTLPL  
NRNIPPIPNRADVSNLILIDKIEFIPIISSVQNRKQKLEITIGKINLPFNHTKN  
TLNEATNYID"

ORIGIN  
Query Match 81.3%; Score 1794.4; DB 6; Length 2085;  
Best Local Similarity 92.7%; Pred. No. 5.5e-234;  
Matches 1908; Conservative 0; Mismatches 141; Indels 9; Gaps 2;  
QY 154 ATGTGTCAAGGAATACACAATATGTGATTAATTTGAGACATTTGCTAGTCTGATACA 213  
DB 1 ATGTGTCAAGGAATACACAATATGTGATTAATTTGAGACATTTGCTAGTCTGATACA 60  
QY 214 ATGTGTCAAGGAATACACAATATGTGATTAATTTGAGACATTTGCTAGTCTGATACA 273  
DB 61 ATGTGTCAAGGAATACACAATATGTGATTAATTTGAGACATTTGCTAGTCTGATACA 120  
QY 274 CTCACTCTATATCCGACCGATAGAAATATAGTGTCTTAATTAATATCTTTTGGTACC 333  
DB 121 CTCACTCTATATCCGACCGATAGAAATATAGTGTCTTAATTAATATCTTTTGGTACC 180  
QY 334 CTATATCTGTCTTTTGGCCCGGGGAAACAAGACAAAACAGTATGACACAATTTATT 393  
DB 181 CTATATCTGTCTTTTGGCCCGGGGAAACAAGACAAAACAGTATGACACAATTTATT 240  
QY 394 AAAATGGAGAAATTTTGTGTATACACGGTTACAGAAAGCATTAACAGCTTAAGTTA 453  
DB 241 AAAATGGAGAAATTTTGTGTATACACGGTTACAGAAAGCATTAACAGCTTAAGTTA 300  
QY 454 CAAACTTTAGAGATTTAGACAATAATTAACAAGCTATTAATACAGCTTAAGATTTGG 513  
DB 301 CAAACTTTAGAGATTTAGACAATAATTAACAAGCTATTAATACAGCTTAAGATTTGG 360  
QY 514 AGAAATTAATAAAGACTACAAAGCTCTGGAATTACACCATCATCATGATTAACAAGCT 573  
DB 361 AGAAATTAATAAAGACTACAAAGCTCTGGAATTACACCATCATCATGATTAACAAGCT 420  
QY 574 GCCTGACCTCTTAATAATGAGATTTGAGAAATGTTCAAGATTTTATTCGAGAAATACCT 633  
DB 421 GCCTGACCTCTTAATAATGAGATTTGAGAAATGTTCAAGATTTTATTCGAGAAATACCT 480  
QY 634 GGTTCGAACCTTGAACCTTAAACCGCTTATCTACTATTTATGCGCAAGCTGTAAT 693  
DB 481 GGTTCGAACCTTGAACCTTAAACCGCTTATCTACTATTTATGCGCAAGCTGTAAT 540  
QY 694 TTTTCATTAATTTATTTATACAAAGGTGTAATTTGCTATGATGAATGACAGATATA 753  
DB 541 TTTTCATTAATTTATTTATACAAAGGTGTAATTTGCTATGATGAATGACAGATATA 600  
QY 754 CATCTTCACAATTTGAACCTTATGCTGGAACATCGATGATTAATTAACCTTTTAAA 813



```

Db      601 CATCTTCACAAATGAACTTAATGCTGGACATGATGACTATTAATCTTTTAAA 660
Qy      814 GAAATATATACCTAATATATAGTAATATTGTGCAAAATACCTATAGACAGACTAAATAAT 873
Db      661 GAAATATATACCTAATATATAGTAATATTGTGCAAAATACCTATAGACAGACTAAATAAT 720
Qy      874 CTTAGAGACGAAACAAATATGAAATGAGATATATTAATGACTATCGAAGATATATGACC 933
Db      721 CTTGAAACGAACTTAATATGAGATGAGATATATTAATGATTAATCGAAGATATATGACT 780
Qy      934 ATTACTGATATATGATACCACTCTCTCAATTTTCTTATATGATATATTAAGATATATGAGAT 993
Db      781 ATTACTGATATATGATACCACTCTCTCAATTTTCTTATATGATATATTAAGATATATGAGAT 840
Qy      994 TCAATAGGAGGAATGAGATTAAGAAAGCATTTAAGAACTGACAAAGAAATTTATACA 1053
Db      841 TCAATAGGAGGAATGAGATTAAGAAAGCATTTAAGAACTGACAAAGAAATTTATACA 894
Qy      1054 ACTGAAATTAATTTTGTATGCTCTCTCAATTTTGTATGAGATGCAACCAATCTAGCTAGATG 1113
Db      895 ACTGAAATTAATTTTGTATGCTCTCTCAATTTTGTATGAGATGCAACCAATCTAGCTAGATG 954
Qy      1114 GATATTAATTTTACGCTGCAAGTTTAAATTTTCAATTTTGAACAATTAATTTT 1173
Db      955 GATATTAATTTTACGCTGCAAGTTTAAATTTTCAATTTTGAACAATTTATATTT 1014
Qy      1174 TATACAGAAATTAATTTTGGGAAATGCTTATGTTGATTTCTAATCTGATGACACT 1233
Db      1015 TATACAGAAATTAATTTTGGGAAATGCTTATGTTGATTTGATTTGCAATCTGATATATGACT 1074
Qy      1234 ACTTATAGCAATTAATTTTGTATGCAAACTTATATGAGAAAGAAAGAGTTTCAACCAACA 1293
Db      1075 ACTTATGCTACGACGAGAACTGAAATTTATATGAGAAAGAAAGAGTTTCAACCAACA 1134
Qy      1294 AAAACAAATTAAGCACTTTGATCTTATTAAGTTTCAATTTGATGATGACATGACCT 1353
Db      1135 AAAACAAATTAAGCACTTTGATCTTATTAAGTTTCAATTTGATGATGACATGACCT 1194
Qy      1354 CCGTATCCCTCAATC--AACCACTTTTATTAATTAATGAACTTAATTTAAT 1410
Db      1195 CCGTATCCCTCAATC--AACCACTTTTATTAATTAATGAACTTAATTTAAT 1254
Qy      1411 GGCTCATCTAAACACACACTCAATATTTGACGAGAGGGCTTTTATCTAATATCAAAAC 1470
Db      1255 AATTCACCTAAGTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1314
Qy      1471 ACAACTTTTCTCAATTTCTAGAAAAGAGCTGCAATCTAATTTATGATCAGGTTGT 1530
Db      1315 ACAACTTTTCTCAATTTCTAGAAAAGAGCTGCAATCTAATTTATTAATTTAATTTGT 1374
Qy      1531 TCACCAACTTTAATTAATTAATTAATTAATTTTCCATTTTCTAATTTATTTTCC 1590
Db      1375 TTACCAACTTTAATTAATTAATTAATTAATTTTCCATTTTCTAATTTATTTTCC 1434
Qy      1591 TATGATATGATTAACACTCAATATTTAGATACAGGATATTTAGATGACACAGT 1650
Db      1435 TATGATATGATTAACACTCAATATTTAGATATTTATTAATTAATTAATTAATTAATTA 1494
Qy      1651 AGGTGTAATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1710
Db      1495 AGGTGTAATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1554
Qy      1711 GGTAACATCTTGAACAACTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1770
Db      1555 GGTAACATCTTGAACAACTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1614
Qy      1771 TTGGTTATTTTACAAAGTCAAGGGCGTTAGAAATTTACATGTAAGAACTCTTAATTTACA 1830
Db      1615 TTGGTTATTTTACAAAGTCAAGGGCGTTAGAAATTTACATGTAAGAACTCTTAATTTACA 1674
Qy      1831 CAATTTATTTTCAATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1890
Db      1675 CAATTTATTTTCAATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1734

```

```

Qy      1891 ATATCTTTACAAATACAGAGATATAGAAATACCACCTCAACGACTCAACAACATTTT 1950
Db      1735 ATATCTTTACAAATACAGAGATATAGAAATACCACCTCAACGACTCAACAACATTTT 1794
Qy      1951 TCTGGTACAAATTAATTAATTAATTTTACATTAACGAGATTTTGGTATTTTCCATTTCCAGT 2010
Db      1795 TCTGGTACAAATTAATTAATTAATTTTACATTAACGAGATTTTGGTATTTTCCATTTCCAGT 1854
Qy      2011 ACAGTAACATTAACCTTTAATCGAAACATACCATTTATTAATTAATCGGAGATATATCA 2070
Db      1855 ACAGTAACATTAACCTTTAATCGAAACATACCATTTATTAATTAATCGGAGATATATCA 1914
Qy      2071 AATTCATTTTATCATTTGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2130
Db      1915 AATTCATTTTATCATTTGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1974
Qy      2131 AATGAGAAAAACAAATTAAGAACTATTCGAAACAAATTAATTAATTAATTTTCACAAT 2190
Db      1975 AATGAGAAAAACAAATTAAGAACTATTCGAAACAAATTAATTAATTAATTTTCACAAT 2034
Qy      2191 CATACAAAAACACTTTGA 2208
Db      2035 CATACAAAAACACTTTGA 2052

RESULT 4
LOCUS      108083 4451 bp DNA linear PAT 02-DEC-1994
DEFINITION Sequence 1 from Patent EP 0296870.
ACCESSION 108083
VERSION 108083.1 GI:589204
KEYWORDS
SOURCE      Unknown.
ORGANISM    Unknown.
REFERENCE   1 (bases 1 to 4451)
AUTHORS     Ellar,D.J. and Ward,E.S.
TITLE       New toxin-encoding DNA fragments from Bacillus thuringiensis
JOURNAL     subsp. israelensis
            Patent: EP 0296870-A1 1 28-DEC-1988;
            Location/Qualifiers
FEATURES
            source
            1..4451
            /organism="Unknown"
            /mol_type="unassigned DNA"

ORIGIN
Query Match 15.7%; Score 346.6; DB 6; Length 4451;
Best Local Similarity 53.2%; Pred. No. 5.8e-38;
Matches 1176; Conservative 0; Mismatches 934; Indels 99; Gaps 17;

Qy      4 AATCAAAATTAAGATATATGAAATATGATTTGATTTGATGATGATGATGATGATGATGAT 63
Db      888 AATATGAACTCTTATCAAAATTAAGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAA 947
Qy      64 CCGAAGCAAGCAAGTAATGATTTAGATACCTTACACAAATTAATTCAAATCAACATTA 123
Db      948 TTAATATATCTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1007
Qy      124 CAAAACAAATTAACAAAGATGCTCAATATGCTCAAGGAAATATACAAATATGATGAT 183
Db      1008 CAAAGTACAAATTAATTAAGATTTGGCTCAATATGCTCAAGGAAATATACAAATATGATGAT 1067
Qy      184 AATTTGAGACATTTGCTGATGCTGATCAATTTGCTGAGATTTGCTGAGATTTGCTGAT 243
Db      1068 GATTTTGAACCTTTTATTTGAT-----AGTGTGAACCTGAGCTTATATATCTATTTGA 1118
Qy      244 TCCGTAATCTGTTAGCCGATATAGTGGGCTCACTTATATATCCGACCGATAGAAATTA 303
Db      1119 GTTGGACCGTACGATCGGT-----TTGGGTTCAACAACCC 1157
Qy      304 ATAGGTCTAATTAATTAATTTTGGTACCTTAATCACTGTTTGGCCCGGAGAA 363

```

Db 1158 TTAGACTTGGCTTATATAGTTTGGTACATTAATACAGATTCCTTTTCCAGCCCAAGAC 1217  
Qy 364 CAGACAAAACAGATGACACAAATTTATTAATGAGAAATTTTGTGATACACG 423  
Db 1218 CAATCTAACAC--ATGAGTGAAGCTTATTAACAAACTAAATAATATATATAAAAAGAA 1274  
Qy 424 TTAAACA-GAAGCATTAACAGCTTAAAGTTACAACTTTAGAGATTTAGCAAAATTT 482  
Db 1275 ATAGATCAACATATATATAGTAATGCTATATAAATTTTAAACAGTCTGTTATATGTTATC 1334  
Qy 483 ACAAGCTATATATACAGCATATAGATTTGAGAAAATTTAAAGACTACAGCTCCCTG 542  
Db 1335 AGCACTTATCTATATACCTTTAAACATGAGAAATATCCAAACCCCAAAATATCTAG 1394  
Qy 543 ATTACCAACCATCATCAGCATTAACAACAAGCTGCTTGACTCTTAATAATACATTTGAGAA 602  
Db 1395 GATGTAAGGA----CACAATTCAGCTAGTTCAATTAACATTTTCCAAATATGTCATTCGA 1450  
Qy 603 TGTTCACAAATGATTTTATTCGAGAAATACCTGCTTCCAACTTGAACTTTTAAACGCT 662  
Db 1451 GCTTGTAACCTGTCCTCTTAATCTAGTGAATTC-----GATTACTATTAACATCT 1504  
Qy 663 ATTACTACCTATTAATGCGCAAGCTGCTAATTTTCAATTTATTTATTAACAACAAGTGC 722  
Db 1505 AGTATTAATCTAGTATGCAACAGCAGCAACTTACATCTGATGTTATTAATCAAGCCGT 1564  
Qy 723 TGAATTTGCTGATGATGAAATGACAGATATACATCTTACAAATTTGAACCTTAATGCTG 782  
Db 1565 CAATTTGAAAGCGATTTTAAAAACAATGCAATGATTAATTTAGAGCTTTGGC--- 1621  
Qy 783 AACATCAGATGACTTATTAATTAACCTTTTAAAGAAATATACCTTAATATAGTAATTTG 842  
Db 1622 AACGCAATTTATTAATTAATCAAGTATGACTTAAGCTATGAAAGATTAACCTTAATTTG 1681  
Qy 843 TGCAATATCTATTTGAACAGGACTTAAATAATCTTAGAGACGAAACCAATTAAGAAATGAG 902  
Db 1682 TGTATCAACTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTTGA 1741  
Qy 903 TATATTAATGACTATCGAAGATATATGACCATTACTGATTAATGATCCATCTCTCAAT 962  
Db 1742 TGGAAATATTAACCTGGAACACATA-----CAATGATATGAAACAAATAATGACTAGC 1795  
Qy 963 TTTCTTATATGATATTAATAAAGATATAGATTTCAATAGAGAAATAGAAATGAAAGCAT 1022  
Db 1796 TGTATATGATGTTGTCATCTTCTCTAATTAATGATGATGATTAATTCAAATAGGCT 1855  
Qy 1023 TAAGAAATGACTCAACAAGAAATTTATATACAAGTAAATTAATTTGATGCTTCTCA 1082  
Db 1856 CCAATCTGAATCTTACTCGAAGAAATTTAT--CAAGTACTTAATCTCGAAGAAAGCCCTAT 1913  
Qy 1083 ACTTAGAGTTCAACCCCAATCTAGCTACAGATGAAATATTAATTTAAACAGCTGCAAGTTTAA 1142  
Db 1914 AAATATATGACTTTCATATATCAAGAGATTCATTAACGTAAGCC--GCATTTATTTA 1971  
Qy 1143 ATTATTTTCAATTTTGAACAATTTATTTTATATACAGAAATATACAAATTTTGGGAATCG 1202  
Db 1972 CTGGCTTGAATCTTTGAA-----TTTATGAAAAAAGCGAAACTCTCTTAATTAATTT 2026  
Qy 1203 TTTAGTGTATTTCTATTCGATGACCTACTTATAGCAATCTATATCTATACTGAAACTTT 1262  
Db 2027 TTTTACCGCACTTATATATATGTTTCAATTAACACTTATATATATATCCCAAAATCTAG 2086  
Qy 1263 ATATGAGAAAAAGAACAGGTTCAACCAACAACAAACAAATTAAGCACTTTGAATCTTATTA 1322  
Db 2087 TGTTTTGGAAATCACATATGTAATGTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2146  
Qy 1323 AGTTCAATTTGTAATGTAATGTAATGTAATGTAATGTAATGTAATGTAATGTAATGTAAT 1382  
Db 2147 TTAATATTTTATTAATTAATGTAATGTAATGTAATGTAATGTAATGTAATGTAATTAATTA 2206  
Qy 1383 AATTAATCAATTTGAATCTTATTTTAATGGCTCATCTAACAACAACATCAATATTTGAC 1442  
Db 2207 TATATGTAATTAATGATTTTATTAATGTAATGTAATGTAATGTAATGTAATGTAATGTAAT 2266

Qy 1443 AGGAGGCTCT-----TTATCTAATTTATCAAAACAAACTTTTTCATTTCTAGAA 1496  
Db 2267 AGCAGATCTGGGCAATTAATCTTATGATGTAATTAATTAATTAATTTTCCGGTTACCAATCT 2326  
Qy 1497 AAAAGACTGCAATCTACTTATTTGATCCAGGTTGTCACCAACTTTAATTAATCTTATGCA 1556  
Db 2327 TAAACGAAGAGAAATCAAGAAACCTTACCTTTTCCAAATATGATTAATCTATGCA 2386  
Qy 1557 TATTTATCCCATTTTTCATTTATTTACTTATTTCTATGATGTAATTAACAGTACAAAT 1616  
Db 2387 TATTTATC-----ATTATTAATTAAGTCTTATGATATCCCTGCAACATAT 2429  
Qy 1617 ATTAGATACAGTGTAT--TTAGATGACACACAGTATGTTGATATATATATGCAATA 1674  
Db 2430 AAACTCAAGTATATAGTTTGTGCTTGACACACACTGATGTTGATCTTAAATAATCAAT 2489  
Qy 1675 TCGATTAATTAATTAATCAATGATCCCGCAATCAAAAGTAAACATCTTGATACAACTCT 1734  
Db 2490 TATACATTTTAATTAATCAATCCAAATTCAGCTGTAAAGCAATTCACCTTGGGACTGCTCT 2549  
Qy 1735 AAGGTAATGGAAGACCTGTCATACAGAGAACTTGGTTATTTTAAACAAAGTCAAGG 1794  
Db 2550 AAGGTTGTTCAAGACCTGTCATACAGAGGGAATTAATGATTTCCAAAGATCA---- 2605  
Qy 1795 CGTTTGAATAATCATGTGAAATCTCTAATTTTACACAAATCTTATTTCAATTAATGACTGCA 1854  
Db 2606 --TTCAAAATTAATCATGTCAACACTCAAAATTTTCAACAAATGATTTTATTAAGAAATCGT 2663  
Qy 1855 TATGCTCAAAATGATGCTGGAATAATCTCTTAATTAATCTTACAAATACAGAGATA 1914  
Db 2664 TATGCTCAAAATGGAAGCGCAAAATCTGAGCTGTAAATTTTAATTTATCTTATCTCCAGGGG-- 2721  
Qy 1915 ATAGAAATACCACTCAACGCTCAACAACACTTTTCTGGTATCAAAATTAATTAATTTA 1974  
Db 2722 -TAGCAAACTGGGATATAGGCACTCAACCCACTTTTCTGGTATCAAGTATTAAGCAATTTA 2780  
Qy 1975 CAATACGAGATTTTGGGATTTTCCAAATTTCCAAAGTACAGTACATTAACCTTTAAATGCA 2034  
Db 2781 AAATATTAAGATTTTCAAGTATTTGAAATTTTCTAACAGAGTGAATTTGCTCCAAATCAA 2840  
Qy 2035 AACATACATTTAATTTAATTCGTCAGATGTAAT--CAATTCATTTTAATTAATGAT 2091  
Db 2841 AACATATCTCTGTTAATGTTGAGATGTAATTAATTAATTAATTAATTAATTAATTAAT 2900  
Qy 2092 AAAATGAATTTATACCAATTTACTTCTCTATGACACCAAAATTAAGAAAAAATTA 2151  
Db 2901 AAAATTTGAATTTTCTGCCAATTAATCTCTCTATTAAGAGATTAAGAGAAACAAATTA 2960  
Qy 2152 GAAACTATCCAAACAAATAATTAATTAATTTTTCACAAATCATACAAAA 2200  
Db 2961 GAAACGATACAAACAAATTAATTAATTAATTTATCAATCTATTAATAA 3009  
RESULT 5  
109103  
LOCUS 109103 4451 bp DNA linear PAT 02-DEC-1994  
DEFINITION Sequence 1 from Patent WO 8810305.  
ACCESSION 109103  
VERSION 109103.1 GI:588188  
KEYWORDS  
SOURCE  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 4451)  
AUTHORS Ellar,D.J. and Ward,B.S.  
JOURNAL Patent: WO 8810305-A 1 29-DEC-1988;  
FEATURES  
1. .4451 Location/Qualifiers  
/organism="unknown"  
/mol\_type="unassigned DNA"



Db 2901 AAAATTGAATTCCTGCCATTACTCTGTTCTATTAAAGAGGATAGAGAAAACAAAATTTA 2960

Qy 2152 GAAACTATCCAAACAAAATAATATACATTTTTCACAAATCATACAAAA 2200

Db 2961 GAAACAGTACAAACAAATATATATACATTTTATGCAAAATCCTATAAAA 3009

RESULT 6

LOCUS BTX0D1 3543 bp DNA linear BCT 18-APR-2005

DEFINITION Bacillus thuringiensis gene for 130 kDa delta-endotoxin.

ACCESSION Y00423

VERSION Y00423.1 GI:40351

KEYWORDS delta-endotoxin; endotoxin.

SOURCE Bacillus thuringiensis

ORGANISM Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus; Bacillus cereus group.

REFERENCE 1 (bases 1 to 3543)

AUTHORS Ward, E.S. and Ellar, D.J.

TITLE Nucleotide sequence of a Bacillus thuringiensis var. israelensis gene encoding a 130 kDa delta-endotoxin

JOURNAL Nucleic Acids Res. 15 (17), 7195 (1987)

PUBMED 2821500

2 (bases 1 to 3543)

AUTHORS Ward, E.S.

TITLE Direct Submission

JOURNAL Submitted (11-MAY-1987) Ward E.S., University of Cambridge, Department of Biochemistry, University of Cambridge, Tennis Court Rd., Cambridge CB2 1QM

COMMENT \*strain= var. israelensis;

FEATURES

source

1. .3543

/organism="Bacillus thuringiensis"

/mol\_type="genomic DNA"

/db\_xref="taxon:1428"

1. .3543

/note="unnamed protein product; endotoxin (AA 1-1180)"

/codon\_start=1

/transl\_table=1

/protein\_id="CA66485.1"

/db\_xref="GI:40352"

/db\_xref="GOA:P16480"

/db\_xref="UniProt/Swiss-Prot:P16480"

/translation="MNPYONKREYETLAAQSOKLINSNNYTRYPEIENSFKOLLOSTNY KDMLNKCOOQYGGDEFETIDSGELSAVTIVGVLTGFGFTPLGLALIGFTLLP VLPAODSNTMSDPITOTKNIKKBIATVSNANKLINSFVSIQTHHILKTMEN NPNQNTDVTQIOLVHYHONVIBELVNSCPNPSDCDYNTLVLSVAQANLHL TVLQAVFBAVILKNNROFDYLEPLTADIVYPAVTKAIBDYNYCVTTAKGLNLK TTPSNLDGNINMNTYNTRTKMTADVAALPPNDVGRYPGVOSLRELYQVL NFESPYKYDFOYQOESLTPRPHLFTLDELNFEKAKQTPNPFISHVMFHYTLN NISOKSVFNGHNVTDLKSLGATNIYIFLNLVSLDNKYLNNYSKDFITNG TRILKEVLTAGSGOITVDNKNITGFLPIKRENOGNTLPPTDNTSLISFKISL IPATKQVYTFPATHSVDPKNTITHTLTOIPAAYNLSGTSKYYQGGHGGDL IDFDHFKITQSHSFQSYIRIRIYASNSANTRAVNLSIPVAELGMALETFESG TDYTNLKXKDPQYLEPSPNEKFAPNQNSLVFNRSDVYTTVTLIDIKELPITRSL EDRKQKLETVQOIIINFEANPIKNTLOSLELTDVLDQAAVLVCSISLELPIKSLR LDEVNAKOLSRNVLQNGDESAITLGMTSDNITIOEDDPIFKGHILHSGARDID GTFEPYVIFOKIDSXKLPYTRVLRGVSQKVEILVSRSEIDAIMNVPADLWY LYPTSPDCSGNRCEBSAVPANIGTSMILSCQYDTGKHYVCQDSHORSFTIDTA LDTNENIGVWVFKISSPDYASLDNLEIEGPIDEGALSVRHMEKKNMDQWEAKR SETQOAVDAQOALDALFTNVQDEALQDTLQAOIQAELYVQSIPIVYNDMLSDVRS NMYDIYELDARVAQARLYDIRNIIRKNGDFTQGMVHVTGNADVOQIIDSVSLVLS NMSAGVSNVHLQHNHGVLGVIAKSGPGNGVYTLMDWESNGELTFTSCBEGYITK TVDVPDTRDRIRIEIGTRBSFYBSIELICMNE"

old\_sequence

919

/note="g was a in [1]"

/citation=[2]

922

/note="g was a in [1]"

/citation=[2]

old\_sequence

922

/note="g was a in [1]"

/citation=[2]

ORIGIN

Query Match 15.6%; Score 345.4; DB 1; Length 3543;

Best Local Similarity 53.2%; Pred. No. 9.2e-38;

Matches 1173; Conservative 0; Mismatches 931; Indels 99; Gaps 17;

10 AATPAGATATATACGAATATGAAATATGATGATTCGATACCTCACCCTTATTTTCGAAC 69

Db 4 AATCTTATCAAAATATATATATATATGAAACATTAATATCTTCACAAAATTTAAAT 63

Qy 70 AGAAGCATATAGTATTCATAGATACCTTACACAAATATATCAATCAATCAACAAAC 129

Db 64 ATATCTAT 123

Qy 130 ACAATATCAAGAGTGGCTCATATATGTCAGAGGAATACAAATATGATATATTTTC 189

Db 124 ACAAT 183

Qy 190 GAGCATTTGCTAGTGTATATACAAATTCGACGTTATGTCAGATCTATTTGATCCGGT 249

Db 184 GAAACCTTTATTTGAT-----AGTGGTAACCTAGTCTATATCTTATTTGTTGGG 234

Qy 250 ACTCTGTTAGCCGATATAGGTGGGCTTCATTTATATCCGACGATATGATATAGT 309

Db 235 ACCGTACTGACTGGT-----TTGGGGTTCAACACCCCTTAGA 273

Qy 310 GCTATATATATATCTTTTGGTACCTTATATCATGCTCTTTGGCCCGGGAACAAAG 369

Db 274 CTGCTTTAT 333

Qy 370 AAAACATATATGACACAAATTTTAAATGGAATTTTGTGTATACACGTTATACA 429

Db 334 AACAC--ATGAGATGCTTTATATACAAACCTATATATTTATATATATATATATAT 390

Qy 430 -GAAAGCAT 488

Db 391 TCAACAT 450

Qy 489 CTAT 548

Db 451 TATCAT 510

Qy 549 ACCAT 608

Db 511 AGGA---CAACAATCAAGCTATGTTACTTACATTTTCAAAATGTCATTTCAAGCTTGT 566

Qy 609 CAATGATTTTATTCGAAATATACCTGTGTTTCAACTTGAACCTTATATATATAT 668

Db 567 AAATCTTGTGCTCTTATCTTATCTATGATGTC-----GATTACTATATATATAT 620

Qy 669 AACTATTTATGCGACAGCTGCTATATTTTCAATTTATATATATATATATATATAT 728

Db 621 ATCTAGTATATGCAACACACAACTTACATCTGATCTGTATTAATCAAGCCGTCAAAT 680

Qy 729 GCGATGATATGGAATGACATATATATATATATATATATATATATATATATATAT 788

Db 681 TGAAGCTATTTTAAAAACAATGCAATTCGATTTATTTAGAGCTTTGTC--AACAC 737

Qy 789 AGATGAT 848

Db 738 AATTGAT 797

Qy 849 TACTAT 908

Db 798 AACTAT 857

Qy 909 TAAATGAT 968

Db 858 TATAAATGTAACAT 911

Qy 969 ATATGAT 1028

Db 912 AGATGTTTGTGACCTCTTCTTATATATATATATATATATATATATATATATATAT 971

1029 TGAACCTACAGAGAAATTTATACAGTGAATAATTTGATCGTCTCAACTAG 1088  
1088 TGAACCTACAGAGAAATTTATACAGTGAATAATTTGATCGTCTCAACTAG 1088  
972 TGAACCTACAGAGAAATTTATACAGTGAATAATTTGATCGTCTCAACTAG 1029  
1089 AGTTACACCAATCTAGCTAGCAGATGATATATTTAACAGCTGCAAGTTTAAAT 1148  
1089 AGTTACACCAATCTAGCTAGCAGATGATATATTTAACAGCTGCAAGTTTAAAT 1148  
1030 TATGACTTTCATATCAAGAGATTCACCTACAGTACAC--GCATTTATTTACCTGGC 1087  
1149 TTTATTTTGAAGCAATTTATTTTATATACAGAAATTCATTTTGGGATCGTTAGT 1208  
1088 TTTATTTTGAAGCAATTTATTTTATATACAGAAATTCATTTTGGGATCGTTAGT 1142  
1209 TGTATTTTCTATCTGATGACCTACTATATAGCAATCTATATAGCAATCTTATATG 1268  
1143 CAGCATATATATATGTTTATATACCACTGATATATATATCCAAAATCTAGGTTT 1202  
1269 AGAAAGAAAGCTTACCCACCAACAAATTAAGACCTTGAATCTTATTAAGTTTC 1328  
1203 TGGAAATCAATGATGATGATTAATTAATCTCTGGTTGGCAAAATTTTATAT 1262  
1329 AATTGATCTGATAGCAATCACTCCCTGTTCCCTATACGACACTTATATATTA 1388  
1263 TTTTATATTAATGATGATGATGATTAATTAATTAATTAATTAATTAATTAAT 1322  
1389 TCAATTTGAATCTTATTAATGATGATGATGATGATGATGATGATGATGATGAT 1448  
1333 TAAATGATTTTATTAATGATGATGATGATGATGATGATGATGATGATGATGAT 1382  
1449 GTCT-----TATCTAATTAATCAAAACAACTTTTTCATTTCTTGAAGAAAGAA 1502  
1383 ATCTGGCAAAATTAATGATGATGATTAATTAATTAATTAATTAATTAATTAAT 1442  
1503 CTGCAATCTGATTAATGATGATGATGATGATGATGATGATGATGATGATGAT 1562  
1443 AAGAGAAATCAAGAAATCCCTATCTTTTTCATTAATTAATTAATTAATTAAT 1502  
1563 ATCCATTTTCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1622  
1503 ATC-----ATTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1545  
1623 TACAGCTGTA--TATGATGATGATGATGATGATGATGATGATGATGATGAT 1680  
1546 CAAGGTATAGCTGTTGATGATGATGATGATGATGATGATGATGATGATGAT 1605  
1606 CAATTAATCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1665  
1741 ATGAAAGACCTGATCATACAGAGAGAAATCTGTTTATTAATTAATTAATTAAT 1800  
1666 GTTCAAGACCTGATCATACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1719  
1801 GAAATTAATCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1860  
1720 AAAATTAATCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1779  
1861 ACAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1920  
1780 TCAATTAATCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1836  
1921 ATACACCTCAAGCAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1980  
1837 GAACTGGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1896  
1981 GGAATTTTGGTATTTTCAATTTTCAAGTACAGTACATTAATTAATTAATTAAT 2040  
1897 AAAATTTTCAATTTTCAATTTTCAAGTACAGTACATTAATTAATTAATTAAT 1956  
2041 CCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2097  
1957 TCTCTTGTGTTTAACTGTTGATGATGATGATGATGATGATGATGATGATGAT 2016  
2098 GAATTTATCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2157

Db 2017 GAATTTGCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2076  
Qy 2158 ATCCAAACAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2200  
Db 2077 GTACAAACAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2119

RESULT 7  
BACISRH4 4253 bp DNA linear BCT 17-FEB-1998  
LOCUS Bacillus thuringiensis israelensis plasmid gene for 130 kDa  
DEFINITION insecticidal protein (ISR4), complete cds.  
ACCESSION D00248.1 GI:216289  
VERSION D00248.1  
KEYWORDS 130 kDa insecticidal protein (ISR4),  
SOURCE Bacillus thuringiensis serovar israelensis  
ORGANISM Bacillus thuringiensis serovar israelensis  
Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus; Bacillus  
cereus group.  
REFERENCE 1 (bases 1 to 4253)  
AUTHORS Sen, K., Honda, G., Koyama, N., Nishida, M., Neki, A., Sakai, H.,  
Himeno, M. and Komano, T.  
TITLE Cloning and nucleotide sequences of the two 130 kDa insecticidal  
protein genes of Bacillus thuringiensis var. israelensis  
JOURNAL Agric. Biol. Chem. 52, 873-878 (1988)  
COMMENT The genes of ISR3 and ISR4 that were 130 kDa insecticidal  
proteins of BTI were sequenced and compared with other insecticidal  
protein genes. ISR3 and ISR4 were identical in a region of the  
C-terminal 467 amino acids.  
FEATURES  
Source  
1..4253  
Location/Qualifiers  
/organism="Bacillus thuringiensis serovar israelensis"  
/mol\_type="genomic DNA"  
/strain="HD522"  
/sub\_species="israelensis"  
/db\_xref="taxon:1430"  
/clone="pBGH4"  
/plasmid="pBTI-5 and pBTI-6"  
/note="976 bp downstream of HindIII site"  
381..385  
/note="Shine-Dalgarno sequence"  
393..3935  
/codon\_start=1  
/transl\_table=11  
/product="130 kDa insecticidal protein (ISR4)"  
/protein\_id="BAA00179.1"  
/translation="MNPYONKNEYETLNASQKLNISNNRYPIENSFKLOLSTNY  
KDMLMQONQVGGDFETPIDSGELASVTVGVVLGPGPTPLGLALGFTLIP  
VLPRAQDSNTWSPFITQTKIILKEIASVTVGVVLGPGPTPLGLALGFTLIP  
NPNPONTQDVRTQIQLVHYHFOVNIPELVNSCPSPSDCYNNILVSYQAAHL  
TVLAQAVFEAVYLKNNRQFDYLBPLIDYVPLVTKAIDBYNYCVTVYKGLNLK  
TTPDSNLGNINMNTYVRYTKMTAVLADLAPNVYGVGVIGVQSELRYEIVOL  
NFESPXYVDFVQEDSLRPHLFTWDLINFEKQCTPNNEFTSHYMHFTLD  
NISQSVFQGNHNTDLKSLGLATNITFIPLVNLISDNKYLINDNNISKDPFTNG  
TRLLEKELTAGSGQITDYVKNITFGPLIKRRBNQNTLPPTYDNVSHLISKISLS  
IPATIKOVYTPAFTSSVDPKNTITLHITQIPAVKANSIGTASKVQGFHGTGDL  
IDPDHFETIKQHSNFOQSYFIRIRVANSNGANRVAIVNLISPGVALGMLNPFSG  
TDYTNLKKYDFQVLEFSNEVFAPNQNSLVFNRSVDVYNTTVLIDKLEPLITSIR  
EDREKQLETVQOILNFPYANPIKNTLOSLELDVDDAANLVECSISELYPKEMLL  
LDEVKNALQSOSRNVLQNGPESATLGWTSNDITIOEDDPIFGKHLHSGKALD  
GTIPPTIYFOKIDSKLKPTRYLYRGVSSKQVLEIVSYGEIDAIMVPAIDANY  
LPSFTFDEGSNRCBTSANPANTIGTSMLSQCYDTKSKVVCODSHQFSTTDGTA  
LDITENIGVWVWFKISSPDYASLDNLEVEGPIDEGALSIRKMEKMDQMAK  
SETQOADVAAQKIDALFTNVQDABLQDFTLAQIOVAYEVLOSIPYVYNMLSPVG  
MNYDIYELDARVQARLYLDTNRIKNGDFTQVGMGHVGNADVOQIDGVSVLVS  
NMSAGVSONVHLOHNGVYLVIAKKSQPGVYTLMDCENQERLFTSCBEVYTK  
TVDPPTDTRIRIEIGTESFYIESIELICMNE"

Query Match 15.6%; Score 345; DB 1; Length 4253;  
Best Local Similarity 53.2%; Pred. No. 9.7e-38;

Matches 1175; Conservative 0; Mismatches 935; Indels 99; Gaps 17;  
QY 4 AATCAAAATTAACGATATACGAATATGAAATTATGATGCGATACCTACCTATATTT 63  
Db 390 AATATGAAATCCCTTATCAAAATTAATAATGATATGAAACATTAATATGCTTCACAAAAA 449  
QY 64 CCGAAGAAAAAGATATGATTTCTTGATACCTTACACAAATTAATCCAAATGCAACATTA 123  
Db 450 TTAATATATCTAAATTAATTAACAGATATCCAAATGAAAAATAGCCAAAACAATTAATA 509  
QY 124 CAAAACAAATTAACAAAGATGAGCTCAATATGCTCAAGGAAATACAAATATGATGAT 183  
Db 510 CAAAGTCAAAATTAATTAAGATTTGGCTCAATATGCTCAACAAATCAGACATATGATGGA 569  
QY 184 AATTTGAGACATTTGCTAGTCTGATACAAATGCTGACATTAAGTACAGGTACTATTTGA 243  
Db 570 GATTTTGAACCTTTTATTTGAT-----AGTGTGAACCTCAGTGCCTATATCTATTTGTA 620  
QY 244 TCCGGTACTCTGTTAGCCGATATAGGTGGGCTCACTTCTATATCCGACCGATAGAAATA 303  
Db 621 GTTGGACCGTACGACTGT-----TTCCGGTTGCAACAACCC 659  
QY 304 ATAGGTCTATATATATCTTTTGGTACCCCAATCACTGTCTTTTGGCCGGGAGAA 363  
Db 660 TTAGGACTTGGCTTATATGTTTGGTACATTAATACAGATTTCTTTTCCAGCCCAAGAC 719  
QY 364 CAAGCAAAAGATATGACACAATTTATTAATGGAAGAAATTTTGTGATACACG 423  
Db 720 CAATCTAACAC--ATGAGTGAATTTATTAACAACAACTAAAAATATTAATAAAAAA 776  
QY 424 TTAACA-GAAAGCATTAACAGCTAAAGTACAACTTTAGAGATTTAGACAAATTT 482  
Db 777 ATACATCAACATATATATAGTATGCTAATTAATTTTAAACAGTCTTTAATGTTATTC 836  
QY 483 ACAAGCTATATATCAGCATTAGATGATGGAAGAAATTAAGAAAGTACAGACCTCCG 542  
Db 837 AGCATTAATCAATCACTTAATCACTTAATGAGAAATTAATCCAAACCCCAAAATATCTAG 896  
QY 543 ATTACCAACCATCATCAGCATTAACAACAAGCTGCTTGACTCTTAATATACATTTGAGA 602  
Db 897 GATGTAAAGA----CACAAATCCAGCTAGTTCAATTACATTTTCAAAATGTCATTTCCGA 952  
QY 603 TGTTCACATGATTTTATTCAGAAATATCCTGTTTCCACTTGAACCTTAATTAAGCT 662  
Db 953 GCTTGTAAACTCTGTCTCTCTTAATCTTATGATGATTC-----GATTACTTAACATATCT 1006  
QY 663 ATTACTACCTATTTATGCGAAGCTGCTAATTTTCAATTAATTAATTAACAACAAGGCT 722  
Db 1007 AGTATTAATCTAGTTATGCAAGCAAGCAACCTTACATCTGACTATTAATTAACAAGCT 1066  
QY 723 TGAATGCGCTGATGAATGGAATGCGAGATATACATCTTCACAAAATGGAACCTTAATGCTGG 782  
Db 1067 CAAATTTGAAGCGATTTTAAAAAACAATCGAACATTTGATTTAGAGCCTTTGGC--- 1123  
QY 783 AACATCAGATGACTTATTAATTAATTTTAAAAAGAAATATACCTTAATATAGTAATTTG 842  
Db 1124 AACGCAATTAATTAATTAATTCAGATATGACTTAAGCTATAGAAATTAACATTAATTTG 1183  
QY 843 TGCAAATACCTATATGAGACAGGACTAAAAATCTTAGAGACGAACCAATTAAGAAATGAG 902  
Db 1184 TGTATCAACCTTATTAATAAAGATTAATTAATTAATTAACGACGCTGATAGTAATCTTGA 1243  
QY 903 TATATTTATGACTATCGAAGATATATGACCATTAATGATTAATGATACATCTCTCAAT 962  
Db 1244 TGAATAATTAATCGAAGACATA-----CAATACGATGGAACAAAAATGACTACAGC 1297  
QY 963 TTTCTTATATGATATTAATAAAGATATAGATTTCAATAGAGGAATAGAAATGAAAGCAT 1022  
Db 1298 TGTATTAATCTTGTGTGACTCTTCTCAATTAATGATGATAGTAAATATCCAAATGAGGT 1357  
QY 1023 TAAAGATGACCTCAAGAAGAAATTTATACAACTGAATTAATTTGATGCTTCTCA 1082  
Db 1358 CCAATCTGAATCTACTCGAATAATTTAT--CAGGTACTTAATCTTGAAGAAAGCCCTAT 1415

QY 1083 ACTTAGATTCAAACCAATCTAGCTACATGGAATATATTTAAACAGCTGCAAGTTTAA 1142  
Db 1416 AAATATATATGACTTTCATATATCAAGAGATTCATCTACAGTAGACC--GCATTTATTTA 1473  
QY 1143 ATTAATTTCAATTTTAAACAATTTATTTTATATACGAAAAATACAAATTTGCGGAATCG 1202  
Db 1474 CTGGCTGTGATCTTTTAA-----TTTTTATGAAAAAGCCCAACCTCTCTAATTAATTT 1528  
QY 1203 TTTAGTTGATTTTCTAATCGATGACACTTATAGCAATACATTAATCTGAAACCTTT 1262  
Db 1529 TTTCAACAGCATTATATATATGTTTCAATTAACACTTGATATATATATCCAAAAATCTAG 1588  
QY 1263 ATATGAGAAAGAACAGGTTCAACCAACAACAAAACAATAAGACATTTGAATCTTAATA 1322  
Db 1589 TGTTTTGGAAATCACATGTAACTGATTAATTAATTAATCTTGGTTGGCAACAAATAT 1648  
QY 1323 AGTTCAATTTGTAATGATATGACATACCTCTGTTTCCCTATTTACCAACACTTAT 1382  
Db 1649 TTAATATTTTATTAATATGTCATAGCTTAGATTAATTAATATCTAATATGATTAATA 1708  
QY 1383 AATTAATCAATTTGAACCTTATTTAATGCTCATCTTAACAACACACTCAAAATATTCAGC 1442  
Db 1709 TATTAATGAATGATTTTATTAATTAATATGATATGATCTTGGAGAAAGAACTTAC 1768  
QY 1443 AGAGGGTCT-----TTATCTAATTAATCAAAACAACTTTTTCATTTCTAGAAA 1496  
Db 1769 AGCAGATCTGGGCAATTAATCTTATGATGTAATTAATAATTTTCCGGTTACCAATTC 1828  
QY 1497 AAAAGACTGCAATCTATTTATGATCCAGGTTGTTCACAACTTTAATACTATAGTCA 1556  
Db 1829 TAAACGAAGAGAAATCAAGAAACCTTACCTTTTTCAAACATATATTAATCTATATCTCA 1888  
QY 1557 TATTTATCCATTTTTCATTAATTAATTAATTTCTATGATGATTTAGATTAACGTACAAT 1616  
Db 1889 TATTTTATC-----ATTATTAAGTCTTATGATACCTCGCAACATAT 1931  
QY 1617 ATTAGATACAGGTGA--TTAGATGGAACAACAGTATGTTGATATGATATATATGCAATA 1674  
Db 1932 AAAATCTAAGGTATATGTTTGGCTTGGACACACTGTATGTTGATCTTAATAATACATTT 1991  
QY 1675 TCGATTAATAATTAATTAACATGATCCCGCAATCAAAAGSTAAACATCTTGTATCAAACTCT 1734  
Db 1992 TATACATTAATTAATCTCAAAATTCAGCTGTAAACGAATTAATCACTTGGACCTGCTCT 2051  
QY 1735 AAGSTAAATGAAGACCTGCTCATACAGAGAGAACTTGGTTATTTTAAAGTCAAGG 1794  
Db 2052 AAGGTGTTCAAGGACCTGCTCATACAGAGAGGGAATTAATGATTTCAAAAGATCA---- 2107  
QY 1795 CGTTAGAAATTAACATGTGAACCTCTTAATTTTACACATCTTATTTTCAATTAAGCTTGA 1854  
Db 2108 --TTTCAAAATTCATGTCAACACTCAAAATTTTCAACATCGTATTTTATAGAAATTCGT 2165  
QY 1855 TATGCTCAAAATGCTGTGAATAATCTCTTAATATCTCTTAACAATACAGAGAGTA 1914  
Db 2166 TATGCTCAAAATGGAAGCGCAAAATACAGAGCTGTATTAATCTTAATATCTTCCAGGGG-- 2223  
QY 1915 ATAGGAATTAACACTCAAGCACTCAACAACATTTTCTGGTACAAATTAATTAATTTA 1974  
Db 2224 -TAGCAAACTGGGTATAGGCACTCAACCCACTTTTCTGTACAGATTAATAGAAATTTA 2282  
QY 1975 CAATACGGAATTTTGGGTATTTTCAATTTTCAAGTACAGTAACTTAATTAATCGA 2034  
Db 2283 AAATATTAAGATTTTTCAGTACTTAATTTTCTTAACAGAGTGAATTTGCTCCAAATCAA 2342  
QY 2035 AACATACATTAATTAATTAATCGTACAGATAT--CAATTCATTTTAATATGAT 2091  
Db 2343 AACATATCTCTGTGTATATGTTTGGATGTATATACAAACAAGATCTTATTTGAT 2402  
QY 2092 AAAATGAATTTATACCAATTAATCTCTATGCAACCAAAATAGAGAAAAAATA 2151  
Db 2403 AAAATGAATTTTCTGCCAATTAATCTGTTCTATTAAGAGAGATAGAGAAAAAATA 2462





```

Oy 1443 AGGAGGCTCT-----TTATCTAATTATCAAAACAACTTTTTCATTTCCTAGAAA 1496
Db 2745 AGCGAGGATCTGGGCAAAATTAAGTAAATTAATAATTTTTCGGGTATCCAAATTCCT 2804
Oy 1497 AAAAGACTGCATCTAGTTATTTGATCCAGGTGTTCCACCAACTTAAATCTATAGTCA 1556
Db 2805 TAAACGAAGAGAAATCAAGAAACCTTATCCACTTTTTCACATATGATTAAGTACGA 2864
Oy 1557 TATTTTATCCATTTTTCATTAATTTTCTATGATGATTTGATTTACAGCTACAAAT 1616
Db 2865 TATTTTATC-----ATTTATTAATAAGCTTAGATCCCTGCAACATAT 2907
Oy 1617 ATTAGATACAGGTGA--TTAGATAGACACACAGTATGTTGATAGATATATGCAATA 1674
Db 2908 AAAACTCAAGTATACGTTTCTTGACACACTCTAGTGTGATCCCTAAAAATACAAATT 2967
Oy 1675 TCAGATAAATTAATTAACAATGATCCCAACAACTTGAATCAAACTTATGATCAAACTCT 1724
Db 2968 TATACACATTTAATCACTCCCAATTCAGCTGTAAAGCAATTCACCTGGGACTGCTTCT 3027
Oy 1735 AAGGTAATGAAGAAGCTGTCATACAGAGAGAACTTGTTTATTTACAAAGTCAGAGG 1794
Db 3028 AAGGTGTTCAAGGACTGTCATACAGAGAGGATTTAATTTGATTTCAAGATCA---- 3083
Oy 1795 CGTTAGAAATTAATCAATGAACTCTAATCTTACACAACTTTATTTATTTAGACTTGA 1854
Db 3084 --TTTCAAAATTAATCAATGCAACCTCAAAATTTTCAACATGCTATTTTATTAAGAAATCGT 3141
Oy 1855 TATGCTCAAAATGAGTGGTGAATAATCTCTCTAATATATATCTTACCAATACAGAGATA 1914
Db 3142 TATGCTCAAAATGAGAGGCAAAATCTGAGCTGTATTAATCTTATGATCCAGGG-- 3199
Oy 1915 ATAGAAATACCACTCAACGACTCAACAACTTTTCTGGTACAAATTAATTAATTTA 1974
Db 3200 -TAGAGAACTGGGATAGGACTCAACCCACTTTTCTGGTACAGATTAATACGATTTA 3258
Oy 1975 CAATACGAGATTTTGGGTATTTCCAAATTTCCAAATGACATTAACCTTTAATATGA 2034
Db 3259 AAAATATAAAGATTTTCACTACTTGAATTTTCTAAGAGAGGAAATTTGCTCCAAATCA 3318
Oy 2035 AACATACATTTATTAATGATGAGATGAT---CAAAATCAATTTATCATTTGAT 2091
Db 3319 AACATATCTCTGTTGTTTATGCTTGGATATATATACAAACACACAGTACTTATGAT 3378
Oy 2092 AAAATGATTTATTAACAATTAATCTCTCTATGACCAAAATAGAGAAAACAAAATTA 2151
Db 3379 AAAATGATTTTCTGCAATTAATCTGTTCTATAGAGAGATAGAGAAACAAAATTA 3438
Oy 2152 GAAACTATCCAAACAAAATTAATGATTTTTCACAAATCATACAAAA 2200
Db 3439 GAAACAGTACAACAAAATTAATGATTTTATGCAAAATCCTATAAAA 3487

```

```

RESULT 9
BTEPTOXIS 127923 bp DNA linear BCT 16-APR-2005
LOCUS BTEPTOXIS 127923 bp DNA linear BCT 16-APR-2005
DEFINITION Bacillus thuringiensis subsp. israelensis plasmid pBtoxis.
ACCESSION AL731825
VERSION AL731825.1 GI:21685410
KEYWORDS
SOURCE Bacillus thuringiensis serovar israelensis
Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus; Bacillus
ORGANISM cereus group.
REFERENCE 1 (bases 1 to 127923)
AUTHORS Berry,C., O'Neill,S., Ben-Dov,E., Jones,A.F., Murphy,L., Quail,M.A.,
Holden,M.T., Harris,D., Zaritsky,A. and Parkhill,J.
Complete sequence and organization of ptoxis, the toxin-coding
plasmid of Bacillus thuringiensis subsp. israelensis
JOURNAL Appl. Environ. Microbiol. 68 (10), 5082-5095 (2002)
PUBMED 12324359
REFERENCE 2 (bases 1 to 127923)
AUTHORS Parkhill,J.

```

```

TITLE Direct Submission
JOURNAL Submitted (19-APR-2002) Submitted on behalf of the pBtoxis
sequencing team, Sanger Centre, Wellcome Trust Genome Campus,
Hinxton, Cambridge CB10 1SA E-mail: parkhill@sanger.ac.uk
COMMENT Details of pBtoxis sequencing at the Sanger Centre are available on
the World Wide Web.
(URL, http://www.sanger.ac.uk/Projects/B_thuringiensis/).
FEATURES
source
1..127923
/organism="Bacillus thuringiensis serovar israelensis"
/mol_type="genomic DNA"
/db_xref="taxon:1430"
/note="plasmid pBtoxis"
1145..1149
RBS
1163..2089
gene
1163..2089
CDS
1163..2089
/gene="pBto01"
/note="Similar in part to Bacillus anthracis pxa1-49
TR:Q9X319 (EMBL:AF065404) (227 aa) fasta scores: E():
8.9e-44, 78.48% id in 158 aa"
/codon_start=1
/transl_table=1
/protein_id="CAD30064.1"
/db_xref="GI:21685411"
/db_xref="UniProt/TREMBL:Q8KNX5"
/translation="MSTHLNFSQNFVHGVKFEFDGLFNITKLENERKEPSSALATY
IMHFEQNDGMLPREFOIMDLAKSGISPTTYTGVOCLRLVLEIPVGNATYVE
LVYALVYHNAETTMINTSLSPRLPFLITOTLSLVKLRNDGNMILLETNT
FSKRIQHEHRYKHIEDPEITRKAPFLKEKNRRAKRYQIEYKPIKPAVVDKDK
SSSRITVRROVOVIEIKFNATFSNVCIEQNDKELHPIAKCREAVSRLLHMQQ
ALRKQKENIMTAFROEIVDIEIYLPYKQKON"
2246..3106
/note="18240"
complement(2246..2262)
/rpt_type=INVERTED
2337..3044
/gene="pBto03"
2337..3044
CDS
2337..3044
/gene="pBto03"
/note="Similar to Bacillus thuringiensis insertion element
18240-a protein TR:Q45766 (EMBL:M23740) (235 aa) fasta
scores: E(): 2.2e-92, 99.57% id in 235 aa, and to
Mycobacterium fortuitum, transposase tmp tmp6100 TR:Q49185
(EMBL:X53635) (254 aa) fasta scores: E(): 1.4e-37, 48.05%
id in 231 aa"
/codon_start=1
/transl_table=1
/product="insertion element 18240 protein"
/protein_id="CAD30065.1"
/db_xref="GI:21685412"
/db_xref="GOA:Q8KNX4"
/db_xref="InterPro:IPR001584"
/db_xref="UniProt/TREMBL:Q8KNX4"
/translation="MERKNI FKMCHYQADILMTVRMYLYRNISFRDLVEMERGIS
LSHTTIRWYHGVPELNERIRKHLKKTNSWRVDEYIYIKGNMVLRYAVSBEQNT
LDFTLSKRDAAKACFLKKALASFYTKPRVITVDNKAVPAVALRELKIEKSIPLYGM
PLRYKVIYKIMVIDHDFIKKRIILNMGLGSMQAVMGINGISAMHVVKKGQQLRLAQ
SAQNRKCIHQDLGLTA"
2538..2933
/gene="pBto03"
/note="HMPfam hit to PF00665, Integrase core domain"
3090..3106
/rpt_type=INVERTED
3521..4381
/note="18240"
complement(3521..3535)
/rpt_type=INVERTED
3612..4319
/gene="pBto04"
3612..4319
CDS
/gene="pBto04"

```



```

/note="Similar to Bacillus thuringiensis insertion element
IS240-a protein TR:Q45766 (EMBL:M23740) (235 aa) fasta
scores: E(): 3.5e-91, 99.14% id in 235 aa, and to
Mycobacterium fortuitum, transposase tnpA or tnp6100
TR:Q49165 (EMBL:X53635) (254 aa) fasta scores: E():
1.1e-37, 48.05% id in 231 aa"
/codon_start=1
/transl_table=11
/product="insertion element IS240 protein"
/protein_id="CAD30066.1"
/db_xref="GI:21685413"
/db_xref="GOA:Q8KX55"
/db_xref="InterPro:IPR001584"
/translation="MEKENIFPKMKHYQADMTLWVRLRYNLSFRDLVEMMERGLS
LSHTITMRVHOYGELENERIKRLKFRINDSRRVDEFTYIKSGENMYLYRVAVSEGRT
IDFYLSKKRDAKACPLKALASPHYRPPVITVDGKAQPAVIRLKNKES1PYGM
PRVRKYLVNMTIEDHRTIKKIRIMRLGLKSMQTAVKMGTAGIEMHMYKQGLKLRQ
SAONQNRCHQHLFGTLA"
/misc_feature
/gene="PBT004"
/note="HMPFam hit to PF00665, Integrase core domain"
/size= 4381
/rpt_type=INVERTED
/complement(4634. .5275)
/gene="PBT005"
/complement(4634. .5275)
/gene="PBT005"
/note="Similar to N-terminus of Bacillus anthracis pXol-18
TR:Q9X219 (EMBL:AF065404) (315 aa) fasta scores: E():
4.2e-56, 84.15% id in 183 aa, and to Bacillus halodurans
Bh2364 protein TR:Q9KXCS (EMBL:AP001515) (378 aa) fasta
scores: E(): 1.6e-18, 35.45% id in 189 aa, and weakly to
Lactobacillus delbrueckii integrase/recombinase orf2
TR:Q48538 (EMBL:Z50864) (333 aa) fasta scores: E(): 6.3,
2e.88% id in 90 aa, and to Bacillus thuringiensis
resolvasse tnpI SW:TNRI_BACTU (P10020) (284 aa) fasta
scores: E(): 8.5, 23.88% id in 180 aa"
/codon_start=1
/transl_table=11
/product="Integrase/recombinase family protein"
/protein_id="CAD30067.1"
/db_xref="GI:21685414"
/db_xref="GOA:Q8KXN3"
/db_xref="InterPro:IPR004107"
/translotion="WSKNINLBETLAAATLIEKGRKOSTIKRYADIDPFKMLRA
NKKLLHLSWSELEADYQTYFSELEDRKRSQETRRHIVWLKELMPCLTSPDLGG
INLSIIPDQSLINDNDPITEMERKLLKQVLSFTGILSTQACATPLINDRVCJINLVV
NGSLQELQSVLSNMIHINIFANINWCYVKIKIDTLNRDLIFVQELIVAFHRTSF"
/complement(4985. .5521)
/gene="PBT005"
/note="HMPFam hit to PF02899, Phage integrase, N-terminal
SM-like domain"
SM-like domain"
/complement(5283. .5286)
/complement(5364. .5507)
/gene="PBT006"
/complement(5364. .5507)
/gene="PBT006"
/note="Similar to Bacillus anthracis pXol-17 TR:Q9X218
(EMBL:AF065404) (47 aa) fasta scores: E(): 2.1e-12, 68.08%
id in 47 aa"
/codon_start=1
/transl_table=11
/product="putative integral membrane protein"
/protein_id="CAD30068.1"
/db_xref="GI:21685415"
/db_xref="UniProt/TREMBL:Q8KXN2"
/translotion="MAVLKDSNYIEMNTYNEIGPKGIFWGLIVVPFWSIMLTATFIM
LCK"
/misc_feature
/complement(5370. .5435)
/gene="PBT006"
/note="1 probable transmembrane helix predicted for PBT006

```

	RBS	by T98MM2.0"
	gene	complement (5515.. .5519)
	CDS	/complement (6451.. .8160)
		/gene="pBc007"
		/complement (6451.. .8160)
		/gene="npBc007"
		/note="Similar to Bacillus anthracis pxa01-16 TR_09X2Y7 (EMBL:AF065404) (569 aa)fasta scores: E(): 0, 96.13% id in 569 aa, and to Bacillus thuringiensis pxa016-1like protein TR:CAC50562 (EMBL:AJ96638) (310 aa) fasta scores: E(): 6.5e-122, 99.67% id in 310 aa"
		E(): /codon_start=1
		/transl_table=11
		/product="conserved hypothetical protein"
		/protein_id="CAD30069.1"
		/db_xref="GI:21685416"
		/db_xref="UniProt/TREMBL:O8KRXI1"
		/translation="MGKSPIIONORSLNGKTNAVVTLLKSMQKVAAPEEKVYSVD TYLDLIQHMTTEERLADVEGRVERAKSPSSRANSTLABKGRRKHLYGTNADLE FVLHKKCTFNIGVIPNTILHMWEDYKQSEFPAYIQSOFIALKRLNLHIEN GLDGRYLTKTHFMNEETEEKAMPYVIISVITPKDFKLXIANKLFIDIAQQOHTET TLKRSLDKODERNKTFPGAMRYFLHKYPHOIRAVIBELTALCTGNLPFKICMQ KGVNRKTRVTVLTLSIHSDPISCKEAGEOHNDPPFVKPYARKKFLFEAVLOENIG ESADMKRPHTVLTXYTCCHOLCSIVTGRGLMDWDREGVPTIKTYLLKHLHOTSOYOI LDTAAKBGTYPLAQHITPKRNSDRQAIVFNGLHSYMSLNITKRMFGNALLKOK FAVYVEESYRNVLKYOEETLFRKVAQVGNLAHYLALETMEKEKLVRRKHRTI PSVRWFLEIADIKLPOEKRLVLEIPKQNLLEPMKTERLVRAGVTTTAPOVLHAM Eik"
	RBS	.complement (8168.. .8173)
	RBS	8947.. .8951
	gene	8959.. .10653
	CDS	/gene="pBc009"
		8959.. .10653
		/gene="pBc009"
		/note="Similar to Bacillus anthracis pxa01-14 TR_09X2Y5 (EMBL:AF065404) (564 aa)fasta scores: E(): 1.6e-191,
Query Match	15.6%; Score 345; DB 1; Length 127923;	
Best Local Similarity	53.2%; Pred. No. 2,4e-38;	
Matches 1175; Conservative	0; Mismatches 935; Indels 99; Gaps 17;	
OY	4	AATCAAAAATAAGTATTAAACAATATTGAATATTGATTCGCATACCCTACTATTTTT 63
DB	96531	AATATGATATCTTATCCAAATTAATAAATGATATGAACAATTAATTCCTCCACAAAAA 96472
OY	64	CCGACAGAAACAGTATGATTCCTAGTACCCTTACACAATTAATGCCAATCAACATTA 123
DB	96471	TTAATATATCTAATATTAATTAACAAGATATCCAAATGAATAATGGTCCAACAATTA 96412
OY	124	CAAAACAACAATTACAAGAAGTGCGTCATATATGTCAAGGAAATCACAAATATGTGAT 183
DB	96411	CAAAAGTCAAATTAATAAGATTGGCTCAATATGTGCACAAATCAGACAGTATGTGTA 96352
OY	184	AATTCGAGACATTTGCTAGTGTGTAACAATTCGTCGAGTTTAGTGCAGATCTATTGTA 243
DB	96351	GATTTTAAACTTTTATTTGAT-----AGTGTGAACCTAGTCCCTATCTATTGTA 96301
OY	244	TCCGTAATCTGTATGCCCGGTATAGTGGGCTCATTTATATCCGGAACGATAGGAATA 303
DB	96300	GTTGGAGCCGTACTGATGAT-----TTGGGGTTTCAACAACACC 96262
OY	304	ATAGTGCTTAATTAATATCTTTTGGTACCTTATCATCTGCTTTTGGCCCGCGGAGAA 363
DB	96261	TTAGACTCTTCTTAATATAGTTTGGTACTTAATACAGTTCTTTTTCACAGCCAAAGC 96202
OY	364	CAAGACAAAACAGTATGACACAAATTTAATAATGGAGAAATTTTGTGATACCCG 423
DB	96201	CAATCTAACAC---ATGGAGTGACTTTAATAACAACAATTAATAATTTTAAAAAAGAA 96145
OY	424	TTAACA--GAAGACATATAACAGCTAATAAGTTACAACTTTAGAAAGTTTAGACAATATT 482
DB	96144	ATAGCAATCAACATATATAGTAATAGCTAATTAATTAATTTTAAACAGGTCGTTTAATGTTATTC 96085

QY 483 ACAAGCTATATACAGATTGATGATGGAGAAATTTAAAGACTACAGCTCTCG 542  
 Db 96084 AGCACTTATATCACTTAAACATGGAGATATCCAAACCAACAAATATACAG 96025  
 QY 543 ATTACCACTATCAGATTACAAAGAGCTGCTGATCTGTAAATATGACTTTGAA 602  
 Db 96024 GATGTAAAGG-----CACAAATCAGCTAGTTCATTTACATTTGATTTCCAG 95969  
 QY 603 TGTTCAAGTATTTTATTCGAAATACCTGTGTTCCAGCTTGAATTTAAAGCT 662  
 Db 95968 GCTTGAATCTGTCTCTCTTAATCTAGTATTC-----GATTACTATTAATCTACT 95915  
 QY 663 ATTACTACTTATTTATGCGCAAGCTGCTATTTTCAATTTAAATTTATACCAAGTGC 722  
 Db 95914 AGTATTTCTAGTATGCAAGAGCAAGCAACTTACATCTGATGTATTTAAATCAAGCGT 95855  
 QY 723 TGAATGGCTGATGATGAGATGAGATATACCTCTTCACAAAATGAGAACTTAAAGCTG 782  
 Db 95854 CAATTTGAGAGCTATTTAAACAAATCGCAATTCGATTTATTTAGAGCTTTGCC--- 95798  
 QY 783 AACATCAGATGACTATTTAACTTTTAAAGAAATATACCTTAATATAGTAACTATG 842  
 Db 95797 AACGCACTGATTTATTTATTCAGATTTGACTAAAGCTATAGAGTTTACACTAATTTATG 95738  
 QY 843 TGCAAATTCCTATAGAGAGCTTAAATCTTATAGAGCAACCAATATGAAATGAG 902  
 Db 95737 TGTATCAACTTATTAAGAAAGATTAATTTAAACGAGCGCTGATGATTAATCTTGA 95678  
 QY 903 TATATTAATGACTATGAGATATATGACCATTAAGTATTTAGATTCATCTGCAAT 962  
 Db 95677 TGGAAATATTAATCGGAACACTA-----CAATTCGATGGAACAAATATGACTATGCG 95624  
 QY 963 TTCTTTATATGATTAAGAAATATAGAGATTCATAGAGAAATAGAAATGAAAGCAT 1022  
 Db 95623 TGTATTTGATCTGTGTCAGCTTTCTTAATTTATGATGATGATTAATTCATTAATGAGT 95564  
 QY 1023 TAAAGATGACTACAGAGAAATTTATACACTGAAATTAATTTGATGCTCTCTCA 1082  
 Db 95563 CCAATCTGAGCTTACTCGAGAAATTTAT--CAGGTACTTAATCTTGAAGAAAGCCCTAT 95506  
 QY 1083 ACTTAGAGTTCAACCCAACTAGCTAGAGTGAATATTAATTAACAGTGCAGATTAA 1142  
 Db 95505 AAATATTAATGACTTTCATATCAAGAGATTCATTAACGTAAGC--GCAATTAATTTA 95448  
 QY 1143 ATTAATTTGATTTTGAACAATTTATTTTATACAGAAATACAAATTTGCGAAATCG 1202  
 Db 95447 CTGGCTGATTTCTTTGAA-----TTTTATGAAAAAGCGAACTACTCTTAATTAATTT 95393  
 QY 1203 TTTAGTTGATTTCTAATCGTATGACCTACTTATAGCAATACTATTAATCTGAACTTT 1262  
 Db 95392 TTTCAACGCACTTATATATATGTTTCACTTACACACTGATATATATCCAAAAATCTAG 95333  
 QY 1263 ATATGAGAAAGAAAGAGTTGACCCCAACAAACAAATAGAACATTTGAACTTTAA 1322  
 Db 95332 TGTTTTGGAAATCACAATGTAAGTAAATTAATCTCTGTTTGGCAACAAATAT 95273  
 QY 1323 AGTTTCAATTTGTAATGATAGCAATCACTCTGTTTCCCTATATCAACCACTTAT 1382  
 Db 95272 TTAATTTTATTAATATGATAGCTTATGATTAATTAATATGATTAATTAATTA 95213  
 QY 1383 AATTATCAATTTGACTTTTATTAATGCTCATCTTACACACACTCAATATTCAGC 1442  
 Db 95212 TATTAGTAAATGATTTTATTAATCAATAGTATGTAAGTATTTGAGAAAGAACTTAC 95153  
 QY 1443 AGAGAGGCT-----TTATCAATTTATCAAAACCACTTTTTCATTTCCAGAAA 1496  
 Db 95152 AGCAGATCTGGGCAATTAATCTTATGATGTAATTAATTAATTTTTCGGTTTACCAATTTCT 95093  
 QY 1497 AAAAGACTGCAATCTAGTATTTGATCCAGGTGTTTCAACCAATTTATTAATCTATGCA 1556  
 Db 95092 TAAAGAAAGAGATCAAGAAACCCATCCCTTTTCCAAATATGATTAATGATGCA 95033  
 QY 1557 TATTTATCCATTTTCAATTTATTTACTTATTTCTATGATGATGATTAACAGTACAAAT 1616

Db 95032 TATTTATC-----ATTATTAAGTCTTGTATCTCTGCAACATAT 94990  
 QY 1617 ATTAGATACAGTGTAT--TTAGATGACACAGATGATGATATATATGCAATA 1674  
 Db 94989 AAAATCAAGTATATAGCTTTGCTTGGACACACTTATGTTTATCTTAAATATCAATT 94930  
 QY 1675 TCAGATTAATTAATTAATCAATGATCCGCAATCAAAAGTAAATCTTGTATCAAACTCT 1734  
 Db 94929 TATACACTTTAATCACTCCAAATTTCCAGCTGTATTAAGCAATTCATTTGGAGCTGCTCT 94870  
 QY 1735 AAGTATTTGAAGACCTGCTGATAGAGAGAACTTGTATTTATTAACAAAGTCAAGG 1794  
 Db 94869 AAGTGTTCAGAGCTGTGCTATACAGAGGGGATTTAATGATTAACAAAGTCA---- 94814  
 QY 1795 CGTTAGAATTTACATGTGAAATCTCTAATTTCAACAACTTATTTCAATGACTTGA 1854  
 Db 94813 --TTTCAAAATTAATATCTCAACTCAACTCAATTTTCAACAAATGATTTTAAAGATTCGT 94756  
 QY 1855 TATGCTAACAAATGCTGCGAAATACTCTTCTAATATATCTTACAAATACAGAGATA 1914  
 Db 94755 TATGCTTCAATGGAAGCGCAAAATCTGAGCTGTATTAATCTTATGATATCCAGGGG-- 94698  
 QY 1915 ATAGAAATACCACTGCAAGCACTCAACACACACTTTTCTGCTCAAAATTAATTAATTA 1974  
 Db 94697 -TAGCAGAACTGGATAGGACCTCAACCCCACTTTTCTGTCAGATTAATACGAATTTA 94639  
 QY 1975 CAATACGAGATTTTGGGTATTTTCCAAATTTCCAAATGACATTAATCACTTTAAATGCA 2034  
 Db 94638 AAATTAATTAATTTTCTGATCTTATGAAATTTTCTTACAGAGTGAATTTGCTCCAAATCA 94579  
 QY 2035 AACATACATTTATTTATTTAATCTGCAATGAT--CAATTCATTTAATCAATTTGAT 2091  
 Db 94578 AACATATCTGTGTTTATTTATCTGCGATGATATATCAACACAGATATTTATGAT 94519  
 QY 2092 AAAATGAATTTTATCCAAATTAATCTCTCTATAGCACCAAAATAGAGAAACAAATTA 2151  
 Db 94518 AAAATGAATTTTCTGCAATTAATCTCTCTATAGAGAGATGAGAGAAACAAATTA 94459  
 QY 2152 GAAATCTCCAAACAAATTAATTAATCAATTTTTCACAAATCATCAAAA 2200  
 Db 94458 GAAACGATACACAAATTAATTAATCAATTTTATGCAAAATCTTATTAATA 94410  
 RESULT 10  
 BTH251977  
 LOCUS  
 DEFINITION  
 Bacillus thuringiensis subsp. medellin cry29a gene for Cry29a protein.  
 ACCESSION  
 VERSION  
 AU251977  
 AU251977.1 GI:17977978  
 KEYWORDS  
 cry29a gene; Cry29a protein.  
 SOURCE  
 Bacillus thuringiensis serovar medellin  
 ORGANISM  
 Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus cereus Group.  
 REFERENCE  
 1  
 Delecluse, A. and Orduz, S.  
 Characterization of two new mosquitoicidal toxins, Cry29a and Cry30a, from Bacillus thuringiensis medellin  
 JOURNAL  
 Unpublished  
 REFERENCE  
 2 (bases 1 to 6009)  
 Delecluse, A.  
 Direct Submission  
 JOURNAL  
 Submitted (22-DEC-1999) Delecluse A., Bacteries & Champignons Entomopathogenes, Institut Pasteur, 25, rue du Dr Roux, 75724 Paris Cedex 15, FRANCE  
 FEATURES  
 Location/Qualifiers  
 source  
 1..6009  
 /organism="Bacillus thuringiensis serovar medellin"  
 /mol\_type="genomic DNA"  
 /isolate="163-131"  
 /sub\_species="medellin"





Qy	181	GATAATTTCCAGACATTTGGCTAGTGCCTGATACAAATTCCTGCAGATGATGCGAGTACTATT	240
Db	39235	AATATATCGGGGGAATTTTGGCTAGTTCGAACTATGTTGGAATGTAGTGCAGGATATTATT	39294
Qy	241	GTATCCGTACTCTGTTAGCCGGATATAGTGGGCTCACTCTATATCCGAGCCGATAGGA	300
Db	39295	GTACTAGGAACCTAATGTTAGAGCTTTTGGCCCTT-----	39330
Qy	301	ATAATPAGTGCTATTAATATATCTTTTGGTACCCTATCACTGCTTTTGGCCCGGGGA	360
Db	39331	GCTTAGCTGACAGTATTAATATCTTTTGGGACTTTGTGCCCATCTTTTGGCAAGATCT	39390
Qy	361	GAACAAAGACAAAACAGTATGAGCACAATTTATTAATAATGGAGAAATTTTGTGTATACA	420
Db	39391	GACCCCTG---CAATATGTTTGGACGAAATTTGTTAACATCGAGAGAGGCCCTATACAAAGA	39447
Qy	421	CCGTTAACGAAAGCATTAATAACAGCTAAAGTTACAACTTTAGAGAGATTTAGACAAATA	480
Db	39448	ATGATATAAAACATTAATTAATGTACTAACTT-----CTATCTGTACACCTATTAATAAAAT	39501
Qy	481	TTACAAAGCTATTAATAACAGCTATGATGATTTGGAGAAATTTAAAGACCTACAAAGCTCT	540
Db	39502	CACTGTGATTAATATCAAGAAATTTTGCATTAATAATGGAGCCAGACGTACACAGCTAAT	39561
Qy	541	GGATTACACCATCATCAGCATTAACAACAAGCTGCTTGCCTTAAATACGATTTGAG	600
Db	39562	GCTAAAGCAGTA-CATGATCTCTTACTACTTAAGAACCTATATATGATTAAGATTTTACA	39620
Qy	601	AATGTTCACAATGATTTTATTTCCAGAAATACCTGTTCCACTTGAACTTATTAACG	660
Db	39621	TATGTTAAAA-----ATATCTGATCAACAAAT	39651
Qy	661	CTATTACTACCTATTTATATGGCAAGCTGTAATTTTCATTTAAATTTATTAACAAGGT	720
Db	39652	CCACACTCCCTGATATGCACAAAATGCTACTTGGCACTTGATTTATTAACACGTCT	39711
Qy	721	GCTGAATTTGGCTGATGATGGAATGCAATATACATCTTTCACAAATGAACTATGCT	780
Db	39712	GCTACCTATTAACAATATATGGCTGCAAAATCAAGTAT-----AATCCAGTACTTTC	39765
Qy	781	GGAACATCAGATGCTATTAATACTTTTAAAGAAATATACCTAATATATGTAATCTAT	840
Db	39766	AATTCATCTAATATCTATCTATGAGGCGCTATTTTAAACGTAAATACAAGAATATCTGCTAT	39825
Qy	841	TGTGCAAAATACCTATAGAACACAGACCTTAAAAATCTTATAGACGACCAAAATATGAATG	900
Db	39826	TGTATATCAAAACGTACATGACAGACCTAACTATGATTTAGAACTAATCTAACCCACATG	39885
Qy	901	AGTATATTTATATGCTATGCAAGATATATGACCATTAATCTGATATGATACATCTCTCA	960
Db	39886	AATATGTAATATCTTACCTTATGAAATATCTCTACATCTGTATGTAATCTTATTTGCTATT	39945
Qy	961	TTTTCTTTATATGATATTAATAAAGATATAGATTTCAATAGAGAAATAGAGTAAAGGC	1020
Db	39946	TTTCCAATATATGACCCAGAAAAATATCC-----AATAGGA	39981
Qy	1021	ATTAAAGAAATCTCAACAAGAAATTTATACAACTGAAATTAATTTTATGCTGTCTCT	1080
Db	39982	GTTAAATCTGAACCTTACAGAGAAGTTTATACAAATGTTAA-----	40022
Qy	1081	CAACTTAGAGTTCAAACCAATCTAGCTACGATGGAATATTAATTTAACAGCTGCAAGTTT	1140
Db	40023	-----TTCAATATCACTTTAGAACCTATACGAACCTGAAAAATGATTACTAGAAATCT	40077
Qy	1141	AAATTAATTTTCATTTTATAGAACAAATTTATTTTATACAGAAATACAAAATTTCCGGAA	1200
Db	40078	ACATTATTTACTTGGATTAACAAAGGCGTTTATTAACAAGAAATCC--TGAAGCATT	40134
Qy	1201	CGTTTAGTGGTATTTTCTAATCGATAGCACCTATCTTATAGCAATCTATTAAGTAACT	1260
Db	40135	CTGATCTCTTAATGATATTTTCTTTTAAAGGTAAACAGATGCGCTTTTACACATCTAAT	40194

QY	1261	TTATATGAGAAACAACAGGTTCCCAACAACAAAACATTAAGACATTGGAATCTTAT	1320
Db	40195	GATGATCGCAACATAAATCTGGGGGCGGCTTCAGAAATATATTTTCAAGACATCC	4025
QY	1321	AAAGTTTCAATTTGAACTGATAGACAAATCACTCTCTTTCCTTATTCACACACTT	1380
Db	40255	AAAGTTATTTCTTTTATATAGAAAACAACTATGATTAAGTCCAAATGG--CAGCATTA	40311
QY	1381	ATTAATTAATCAAAATGGAATTTATTAATGGCTCATCTAACAACAACACTCAAAATATTA	1440
Db	40313	GAGAGTACTGAGATATATATATGAATATGATATTTTTTGGAAATAGCAGGAAGTATTC	4037
QY	1441	GCAGAGAGGTCCTTATCTAATTTATCAAAACAACACTTTTTCATTTCTTACAAAAAA	1500
Db	40373	G-----ATATTCATCAATTCMAACATAGAAATATATATTAAGAACT-----	40411
QY	1501	GACTGCAATCTAGTATTTGATCCAGGTTGTTCAACAACCTTATATACTATAGCATATT	1560
Db	40417	-----GATTCATATGATTTCCAAACAAACATGAAAAATTAAGATATGTGCTACT	40477
QY	1561	TTATCCATTTTTCATTATTTACTTATTTCTATGTGATGATTAAGCTACCAATATTTA	1620
Db	40471	CTATCGATATATAAACTGATTAATTTAT-----ATATTTCAAGTTAGAGAAGA	4052
QY	1621	GATACAGGTGATTTAGATGAGACAACAAGTATGTTGATATGATATATGCAATATCAGAT	1680
Db	40522	AGAAAGATGTGATTTATGTTGACACATACATGATGTTGATTTCCAAATATCAATATGATTTA	4058
QY	1681	AAATATATTAACATGATCCAGCAACATCAAAAGTATACATCTTGATACAACTCTAAGSTA	1740
Db	40582	GATATACATCAACCAATTCACAGCTCTTAAAGCTTTGAAGTAAAGTTCTGATTCGAAATTT	4064
QY	1741	ATTGAAGGACCTGCTCATACAGAGAGAAACTTGGTTATTTTACAAAGCTCAAGGCGTTTA	1800
Db	40642	GTGAAGAGTCTGTGTCACACAGGAGGAGACTTGGTATTTCTTAAAGATAGTATGATTTT	4070
QY	1801	GAAATTAATATGTAAGAACTCTTAATTTCTACACATCTTATTTCAATAGACTTGATATGCT	1860
Db	40702	AGAGTTGATTTTTTAAAAAATGTTTCTCGACAAATATCAAGTACGTA---TTGATTAATGCT	4075
QY	1861	ACAATAGTGTGGAATTAATCTCTCTCAATATATATCTTCAATATACAGAGATTAATAGA	1920
Db	40759	ACTATATGCTCAAAAGACACAGT--ATTCTTAAACCGAATAGATATCTATTAAGTGTGAGC	40811
QY	1921	ATACCACTCAACAGCACTCAACAACACTTTTTCTGTATCAAAATTAATATATTTATTAACAATC	1980
Db	40817	TCCTATATACACTTCCCGCAAAAC-----CAATGCTACAGATTTTAACATAT	4086
QY	1981	GGAGATTTTGGTATTTTCCAAATTTCCAAATGACGT-----AACATTACTTTAAT	2031
Db	40867	GCAATATTTGGATATATGATATATTTCCAAAGACAGTTCCAAATATATAAATTTGAAGAGTA	4092
QY	2032	CGAAACATACATTTATATTAATTAATCGGAGAGATATCAAAATCAATTTAATATCATTTGAT	2091
Db	40927	GACACTTATTTAATGACTTATATGTTATGTTACCAATCATTTCAATATATATATATTTGAC	4098
QY	2092	AAATATGAATTTATACCAATTAATCTCTCTATGACACCAAAATAGAGAAAAACAAATTA	2151
Db	40987	AAATATGAATTTATTTCAATCACTCAATCTGATATTTGATATATACAGAGAGCAAAATATA	4104
QY	2152	GAAATCTATCCAAACAAATTAATATCAATTTTTCACAAATATATCAAAAAACACTT	2205
Db	41047	GAATAAAACACAGAAATATGAAATGATTTATTTGTTATATTTAAACAAAGTTCTT	41100

RESULT 12			
LOCUS	BACMSOB	3753 bp	DNA linear BCT 26-APR-1993
DEFINITION	B. thuringiensis insecticidal endotoxin gene, complete cd.		
ACCESSION	M16662		
VERSION	M16662.1	GI:143226	
KEYWORDS	Bacillus thuringiensis		
SOURCE	Bacillus thuringiensis		



```
Db      2351 CTATGTTATATAAAAGTATATAT-----ATATTTTCAGTAGTAGAGAAAAGA 2401
Qy      1651 GATCAGGTTATATAGATGACACACAGTAGTGTGATAGATATATAGCAATATACAGAT 1680
Db      2402 AGAAGAGTTGATTTAGTTGACACATCTAGTGTGATTTCCAAAATACATATATTTA 2461
Qy      1681 AAAATTAATTAATGATCCAGCAATCAAGGTAATCAATCTTGATATACAAACTCTAAGGTA 1740
Db      2462 GATACATACCCCAATTCAGGCTCTAAAGCTTTGAAGGTAAAGTTGATTCGAAATTT 2521
Qy      1741 ATTGAAGAGCTGTGTCATACAGAGAAAAGCTTTGTTATTTACAAAGTCAGAGCGCTTTA 1800
Db      2522 GTGAAGGCTCTGTGTCACACAGGTGAGACTGTGTAATTTCTTAAAGATAGATGATTTT 2581
Qy      1801 GAAATTAATGTAAGAACTCTTAATCTTACAAATCTTATTTATGACTTGATAGCT 1860
Db      2582 AGAGTTAGATTTTAAAAAAGTTTCTGACAAATACAGTACGTA---TTGTTATGCT 2638
Qy      1861 ACAATGTGCTGTAATATCTCTTAATATATCTTTACAAATCCAGAGTAAATAGGA 1920
Db      2639 ACTATGCTCCAAAGACACAGCT--ATTCTTACCGGAATAGATCTATAGTGTGAGC 2696
Qy      1921 ATACCACTCAACGACTGACAAACACTTTTCTGTGACAAATTAATTAATTTACAAATAC 1980
Db      2697 TCCCTAGTACACACTTCCGCGCAAAAC-----CAAAATGTAAGATTTACATAT 2746
Qy      1981 GGAGATTTTGGTATTTTCAATTTCCAAAGTACAGT-----AACATTAACCTTTAAT 2031
Db      2747 GCAGATTTTGGATTAATGTAATTTCCAAAGAACAGTTCCAAAATMAAAATTTGAAGAGAA 2806
Qy      2032 CGAAACATACCAATTAATTAATGATGAGATGATCAAAATCAATTTTAATCAATGAT 2091
Db      2807 GACACTTTATTAATACCTTATATGTAACCAAAATCAATTCATTAATTAATATATTTGAC 2866
Qy      2092 AAAATGTAATTTATCAATTAATCTCTATGACCAAAATAGAGAAAACAAAATTA 2151
Db      2867 AAAATGTAATTTATTCATCAATCACTCAATCTGATTAATTAATGAGAGAGAAAATTA 2226
Qy      2152 GAACTATCCAAACAAAATTAATCAATTTTTCACAAATCAATCAAAACACTT 2205
Db      2927 GAAAAACACAGAAATAGTGAATGATTTATTTGTAATTAATMAACAAAGTTCTT 2980

RESULT 13
LOCUS   DQ078744 434 bp DNA linear BCT 11-JUL-2005
DEFINITION Bacillus thuringiensis strain LDC-9 cry4A insecticidal protein
(cry4A) gene, partial cds.
ACCESSION DQ078744.1 GI:68348788
KEYWORDS
SOURCE   Bacillus thuringiensis
          Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus; Bacillus
          cereus group.
REFERENCE 1 (bases 1 to 434)
AUTHORS Mahalakshmi,A., Sujatha,K. and Shenbagarathai,R.
TITLE    PCR analysis and molecular characterization of cry4A gene of
          indigenous Bacillus thuringiensis LDC-9
JOURNAL  Unpublished
          2 (bases 1 to 434)
          Mahalakshmi,A., Sujatha,K. and Shenbagarathai,R.
          Direct Submission
          Submitted (31-MAY-2005) PG Department of Zoology and Research
          Centre, Lady Doak College, Chinnachockikulam, Madurai, Tamilnadu
          625002, India
FEATURES
         source
         1..434
         /organism="Bacillus thuringiensis"
         /mol_type="genomic DNA"
         /strain="LDC-9"
         /db_xref="taxon:1428"
```

```
gene
CDS
<1..>434
/gene="cry4A"
<1..>434
/gene="cry4A"
/note="cry4A delta endotoxin; contains putative conserved
domain of delta endotoxin"
/codon_start=1
/transl_table=1
/product="cry4A insecticidal protein"
/protein_id="AA96321.1"
/db_xref="GI:68348788"
/translation="SSLNFDQSYFIRIRYVNSGANTRAVINLSIPGVAELGMAINPT
FSGDYTNLAKXKDYLFNSNVEKAPQNSLVFNSDVTANTTVLIDKIEPLPTR
SIRDREROKLETVQIINTFYAMPINKTILQSEBLDVIDDGG"

ORIGIN
Query Match 8.3%; Score 184.2; DB 1; Length 434;
Best Local Similarity 70.7%; Pred. No. 1.5e-15;
Matches 275; Conservative 0; Mismatches 108; Indels 6; Gaps 2;

Qy      1815 AACTCTTAATCTTACAAATCTTATTTTCAATGACTTGATATGCTCAAAATGTCGTGG 1874
Db      3 ATCACTCAATTTTCAACATGATGATTTTATTAAGATTTGTTATGCTTCAATGGAAGCGC 62
Qy      1875 AAATACCTTCTATATATATCTTTACAAATCCAGAGTAATAGAAATACACCTCAACG 1934
Db      63 AAATACCTGAGCTGTATTAATCTTAGATCCAGGGG---TAGCAGAACTGGGTATGGC 119
Qy      1935 ACTCAACAACTTTTCTGTGTAACAATTAATTAATTTACATACGAGATTTGGGTA 1994
Db      120 ACTCAACCCCACTTTTCTGTGTAACAATTAATTAATTAATTAAGATTTTCACTA 179
Qy      1995 TTTCCAATTTCCAGTACAGTAATCAATTAATTAATTAATTAATTAATTAATTA 2054
Db      180 CTAGAAATTTTCTACGAGGTAATTTGCTCCAAATCAAAATCAATCTTGTGTTTAA 239
Qy      2055 TCGTGCAATGAT---CAAAATCAATTTTATCATGATTAATTAATTAATTAATCAAT 2111
Db      240 TCGTTGGAATTAATTAATCAACACAGAGTACTTATGATTAATTAATTTGCGCAAT 229
Qy      2112 TACTTCTCTATGCAACAAAATAGAGAAAACAAAATTAAGAACTATCCAAACAAAT 2171
Db      300 TACTGTTCTATTAAGAGAGATAGAGAAAACAAAATTAAGAAACAGTACAACAATTAAT 359

Qy      2172 AAATCAATTTTCCAAATCAATCAAAA 2200
Db      360 TAATCAATTTATGCAAAATCCTTAATAA 388

RESULT 14
LOCUS   E00614 3756 bp DNA linear PAT 29-SEP-1997
DEFINITION dna encoding a polypeptide having insecticidal activity (BtI
endotoxin).
ACCESSION E00614 GI:2168893
KEYWORDS UP 1986005098-A/1.
SOURCE   Bacillus thuringiensis
          Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus; Bacillus
          cereus group.
REFERENCE 1 (bases 1 to 3756)
AUTHORS Aran,E.U. and Toomaseu,J.P.
TITLE    INSECTICIDAL POLYPEPTIDE
JOURNAL  Patent: JP 1986005098-A 1 10-JAN-1986;
          SHINTORO CORP
OS       Bacillus thuringiensis
PN       JP 1986005098-A/1
PD       10-JAN-1986
PF       21-FEB-1985 JP 1985031647
PR       22-FEB-1984 US 84 582506, 22-JAN-1985 US 85 693556 PI
          ARAN EMU UORUFUIRUDO, TOOMASU JIEI PORTOSUKU PC
          C07K15/04,A01N63/02,C07H21/04,C12N15/00//C12P21/00, (C12N15/00, PC
```



Cl2R1.19);  
CC strandedness: Single;  
CC topology: Linear;  
CC hypothetical: No;  
CC anti-sense: No;  
CC \*source: strain=Iseraelensis ONR-60A;  
CC \*source: clone=BTA4-1;  
FH Location/Qualifiers  
FH  
FT 5'UTR <1..925  
FT CDS 926..2971  
FT /product='BRI endotoxin'  
FT mat\_peptide 926..2968  
FT /product='BRI endotoxin'  
FT 3'UTR 2972..3756  
FT Location/Qualifiers  
source 1..3756  
/organism='Bacillus thuringiensis'  
/mol\_type='genomic DNA'  
/db\_xref='taxon:1428'  
ORIGIN  
Query Match 8.1%; Score 178.2; DB 6; Length 3756;  
Best Local Similarity 50.0%; Pred. No. 4e-15;  
Matches 1108; Conservative 0; Mismatches 923; Indels 186; Gaps 19;  
QY 1 ATGAATCAAAATTAAGATATAGCAATATGAAATTTATGATTCGCATACCTCACTTAT 60  
DB 941 ATGAATCATATCAAAATTAAGATATAGCAATATGAAATTTATGATTCGCATACCTCACTTAT 994  
QY 61 TTTCCGAACAAAACAGTATGATTTCTAGATACCTTACACAAATATCCAAATCAACCA 120  
DB 995 GGTTTAGCAAGTCTAATTAATTAATCTAGATATCCATTAAGCAAAATTAACCAATCAACCA 1054  
QY 121 TTACAAAACAAATTAACAAGAGTGGCTCAATATGTCTCAAGGGAATATACAAATATGCT 180  
DB 1055 CTGAAAAACAGATTTACAAAGATTGGCTCAATGTGTCAAGATATCAACAAATATGSC 1114  
QY 181 GATAATTCAGACATTTGTCTAGTCTGATACAAATTCCTGACAGTTAGTGCAGTACTATT 240  
DB 1115 AATATAGCGGGAAATTTTGCTAGTCTGAAACTATTTGTTGAGTTAGTGCAGTACTATT 1174  
QY 241 GTATCCGGTACTCTGTATGCCGGTATAGTGGGCTCACTTCTATATCCGACCGATAGGA 300  
DB 1175 GTAGTAGACACTATGTTAGAGCTTTGTGCCC-----CT 1210  
QY 301 AATAAGGTGCTATATATATCTTTTGGTACCCCAATCACATCTTTTGGCCGCGGA 360  
DB 1211 GTCTTAGCTGAGTATATATCTTTTGGACTTTGTGCCGATCTTTTGGCAAGATCT 1270  
QY 361 GAACAAAGCAAAACAGTATGACACAAATTTATTAATGGAGAAATTTTGTGATACA 420  
DB 1271 GACCCG---CAATAGTTTGGAGGATTTGTTAAACATCGAGGAGGCCATATCAAGAA 1327  
QY 421 CCGTTAACAGAAAGCATTAACAGCTTAAGTTACAACTTTTGAAGAGTTTGAACAATA 480  
DB 1328 ATAGATTAATAACATATATATGTAATACTT-----CTATCGTAACCTATATAAAAT 1381  
QY 481 TTACAAAGCTATATATAGACATTTAGATTTGAGAAATTTAAAGACTCAAGCTCCT 540  
DB 1382 CAATCTGTAATATATCAAGAAATTTTGCATTAATGGAGCCAGCAGTACACAGCTAT 1441  
QY 541 GGATTTACACATCATCATGACATTAACACAGCTGCTTACTTTAAATATGATTTGAG 600  
DB 1442 GCTAAAGCAGTA-CATGATCTCTTACTACCTTGAACCTATATATGATTAAGATTGA 1500  
QY 601 AATGTTCATGATTTTATTCGAGAAATACCTGTTTCAACTTGAACCTTATTAAGCG 660  
DB 1501 TATGTTAAAA-----ATAATGCTAGCTATCGATA 1531  
QY 661 CTATTTACTACGATTTATGCGAAGCTGTAATTTTCAATTTAAATTTATTAACAAGGT 720  
DB 1532 CCAACACTCCCTGCTATGACAAATAGCTACTGGCACTTGAATTTATTAACATGCT 1591

QY 721 GCTGAATGGCTGATGAAATGCAATATACATCTTCACAAAATTGAACCTAATGCT 780  
DB 1592 GCTACCTATATACATATATATGCTGCAAAATCAAGGTATTAATCCAGTACTTCAATT-- 1649  
QY 781 GGAACATCAATGACTTTTATTAACCTTTTAAAGAAATTTACTTAATATATGACTAT 840  
DB 1650 ----CATCTAATTAATCAATGAGGCTATTTAAACGTAATAATCAAGATATATGACTAT 1705  
QY 841 TGTCAATATACCTATAGAACAGACTAAAAAATCTTATAGACGAAACCAATATGAAATGG 900  
DB 1706 TGTATACAAACGTAATGACAGGACTATGATTTGAATTAATTAATTAACCAACATGG 1765  
QY 901 AGTATATTTAATGACTATGCAAGATATATGACATTAAGTATATGATTAATCTCTCA 960  
DB 1766 AATATGATATATCTTACCGTTTGAATGACCTTAAGTATGATTAATGATTTGCTATT 1825  
QY 961 TTTTCTTATATGATATTAATAAGATATAGATTTCAATAGAGAAATAGAAATTAAGGC 1020  
DB 1826 TTTCCAAATTAATGACCCGAAAAATATCC-----AATAGGA 1861  
QY 1021 ATTAAGAATGAACTCAACAAGAAATTTATCAACTGAATTAATTTTGAATGCTTCTCT 1080  
DB 1862 GTTAAATCTGAACTTATCAGAGAAATTTATACAAATGTTAA----- 1902  
QY 1081 CAATTTAGATTCACCCCAATCTAGCTAGATGGAATATATTAACACGTGCAAGTTT 1140  
DB 1903 ----TTCAGATACATTTAGAACCATTAACAGAACTGAATATGATTAAGTAAATCT 1957  
QY 1141 AATATTTTCAATTTTATAGAACAAATTTATTTTATACAGAAATTAAGAAATTTGCGAAT 1200  
DB 1958 ACATTTATTTACTTGGAATTAACCAAGGCGTTTATACAGAAATTC--TCGACACTT 2014  
QY 1201 CGTTAGTGTATATTTCTAATCTGATGACACTTATATGACATATCTATTAAGTAACT 1260  
DB 2015 CTGATACCTTATGATATTTTCTTTTACAGGTACAGATAGGCTTTACACATACATAT 2074  
QY 1261 TTATATGAGAAAGAACAGGTTCAACCAACAACAAATTAAGACATTTGAATCTTAT 1320  
DB 2075 GATGATGCAACATTAATCTGGCGAGCGTTTCAATGAAATATTTATTTCTCAAGACATCC 2134  
QY 1321 AAGTTTCAATTTGAACTGATAGACATCACTCTGTTTCCCTATTTCAACACACTTT 1380  
DB 2135 AAGATATTTCTCTTTTATAGAAACAACTATATGATTAAGTGAATTTG--CAGACATA 2192  
QY 1381 ATATATATCAAAATGAACTTATTTAATGGCTCATCTAACAACACACTCAATATTTCA 1440  
DB 2193 GAGATCTCAGATATATATATGAATGATTTTGTGAATAGCAGTGAATATTT- 2251  
QY 1441 GCAGGAGGCTTTATCTAATTTATCAAAACAACTTTTTCATTTCTAGAAAAAAA 1500  
DB 2252 ----CGATATTCATCAATTCACAAATGAAATAATATTAATAAGAACT----- 2296  
QY 1501 GACTGCAATCAATGATTAATCCAGGTTTCCACCAACTTTAATACTAATGCTATTT 1560  
DB 2297 ----GATTCCTATATGATTTCAAAACAAACATGAGAAATGAAATATATGCTATCT 2350  
QY 1561 TTATCCATTTTTCATTTATTTATCTTATCTATGATTTGATTTGATTTGATTTGATTT 1620  
DB 2351 CTATCTATATTAATAACGATATTAAT-----ATATTTCAAGATTTAGGAAGAA 2401  
QY 1621 GATACAGGTATTTAGATGACACACAGTATGTTGATATGATTAATGCAATATTCAGAT 1680  
DB 2402 AGAAGATGCTATTTAGTTGACACATATGATTTGATTTTCCAAATATCAATGATTTTA 2461  
QY 1681 AAAATATTAATGATATCCAGCAATCAAGGTAACTCTGATTAACAACTCTAAAGTA 1740  
DB 2462 GATTAACATCAACCAATTCACGCTTAATAAACTTTAAAGTAAGTTTAAATTTCAAAAT 2521  
QY 1741 ATTGAAGACCTGGTCACTACAGAGAGAACTGGTTATTTATTTCAAAAGTCAAGGCGTTTA 1800  
DB 2522 GTGAAGGCTCGGTCAACAGTGAATACTGGTATTTCTTAAGATATGATGAATTTT 2581



QY 1801 GAAATTCATGTGAAATCTCTTAATTCACAACTCTTAATTCATTAAGACTTCGATATGCT 1860  
DB 2582 AGAGTTAGATTTTAAAAAATGTTCTCAACAAATCAAGTACGTA---TTGCTATGCT 2638  
QY 1861 ACAAATGCTGTGAAATATCTCTCTTAATTAATCTTCAATTAACGAGTAATAGAA 1920  
DB 2639 ACTAATGCTCAAAAGACAAAGTA---TTCTTAACCGGAATAGTAATCTAATGTCGAGC 2696  
QY 1921 ATACCACTCAACGACTCAACAAACACTTTTCTGTGTAACAAATTAATTAATTAACATAC 1980  
DB 2697 TCCCTAGTACCACTTCGCCGCAAAAC-----CCTAATGCTACGATTTTAATATAT 2746  
QY 1981 GGAGATTTGGTATTTTCCAAATTTCCAAATAGTACGT---AACATTCCTTTAAATCGAA 2036  
DB 2747 GCAGATTTTGGATATGTAACATTTTCAAGAACAGTTTCAAAATTAACATTTGAAGAGAA 2806  
QY 2037 CATACCATTTAATTTA-----ATGTCGAGATGTAATCAATTTCAATTTTAATCAT 2088  
DB 2807 GACACTTTAATTAATGACCTTATATATGTAACCAAAATCAATTCATTAATTAATATATAT 2866  
QY 2089 GATAAAATTTGAATTAATCAATTAATCTCTCTATGCAACCAAAATAGAGAAAAACAAA 2148  
DB 2867 GACAAATTAATTAATTAATTAATCAATCAATCTGATTAATTAATTAATTAATTAATTAAT 2926  
QY 2149 TTAGAACTATCAAAACAAATTAATTAATCAATTTTCAAAATCAATCAAAACACTT 2205  
DB 2927 ATAGAAAAACACGAAATATGTAATGATTTATTTGTTAATTAACAAAGTTCTT 2983

RESULT 15  
LOCUS 103578 1993 bp ss-DNA linear PAT 21-MAY-1993  
DEFINITION Sequence 1 from Patent US 4652628.  
ACCESSION 103578  
VERSION 103578.1 GI:268658  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 1993)  
AUTHORS Walfield,A.M. and Pollock,T.J.  
TITLE Methods and compositions for expression of BTI endotoxin  
JOURNAL Patent: US 4652628-A 1 24-MAR-1987;  
SYNOPSIS Syntro Corporation; San Diego, CA  
FEATURES  
Source location/Qualifiers  
1. 1993  
/organism="unknown"  
/mol\_type="unassigned DNA"

ORIGIN  
Query Match 7.8%; Score 171.6; DB 6; Length 1993;  
Best Local Similarity 53.4%; Pred. No. 4,1e-14;  
Matches 571; Conservative 0; Mismatches 424; Indels 75; Gaps 7;

QY 1 ATGAATCAAAATTAACGATTAATTAACGATTAATTAATTAATTAATTAATTAATTAATTAAT 60  
DB 941 ATGAATCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 994  
QY 61 TTTCCGAAACAGAAACGTAATGATTTCTAGATACCTTACACAAATTAATCAATCAACCA 120  
DB 995 GGTTTTACAGAGTCTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1054  
QY 121 TTACAAACACAAATTAACAGAGTGTCTCAATATGTCAAAGGAAATCAATCAATATGAT 180  
DB 1055 CTGAAAAACAGAAATTAACAGAGTGTCTCAATATGTCAAAGGAAATCAATCAATATGAT 1114  
QY 181 GATAATTTGAGACATTTGCTAGTCTGATTAATTAATTAATTAATTAATTAATTAATTAAT 240  
DB 1115 AATATATGCGGGAATTTGCTAGTCTGATTAATTAATTAATTAATTAATTAATTAATTAAT 1174  
QY 241 GTATTCGGTACTCTGTATGCGGATATAGTGTGCTCACTTCTAATATCCGACCATTAAGA 300  
DB 1175 GTAGTAGGAATCTATGTTAGAGCTTTTGTGCTGCTG----- 1211

QY 301 ATAAATAGTGTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 360  
DB 1212 -TCTTAGTGTAGGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1270  
QY 361 GAACAGACAAACAGTATGACACATTTTAATTAATTAATTAATTAATTAATTAATTAATTAAT 420  
DB 1271 GACCTGTC---AAATGTTTGGCAGATTTTGAACATCGGAGGAGGCTTATACMAAA 1327  
QY 421 CCGTTAACAGAACGATTAACAGTAAAGTTAACAACTTTAGAGATTTAGCAAAATA 480  
DB 1328 ATGATTAACAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1381  
QY 481 TTACAAAGCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 540  
DB 1382 CACTGTGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1441  
QY 541 GGATTTACCAACATCAATCAATTAACAAACAGCTGCTTAAATTAATTAATTAATTAATTAAT 600  
DB 1442 GCTAAGACAGTA-CATGATCTCTTTAATCACTTGAACCTTAATTAATTAATTAATTAATTA 1500  
QY 601 AATGTCACATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 660  
DB 1501 TATGTTAAAA-----ATAATGCTAGCTATGCAATA 1531  
QY 661 CTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 720  
DB 1532 CCAACATCTCTCGATATGACAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1591  
QY 721 GCTGAATTTGCTGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 780  
DB 1592 GCTACCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1648  
QY 781 GGAATCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 840  
DB 1649 ---TCATCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1705  
QY 841 TGTCAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 900  
DB 1706 TGTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1765  
QY 901 AGTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 960  
DB 1766 AATATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1825  
QY 961 TTTTCTTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1020  
DB 1826 TTTCCAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1885  
QY 1021 ATTAAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1070  
DB 1886 GTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1935

Search completed: December 11, 2005, 17:39:58  
Job time : 1160 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OW protein - nucleic search, using frame\_plus.p2n model

Run on: January 20, 2006, 17:46:37 ; Search time 300 Seconds  
(without alignments)  
2019.463 Million cell updates/sec

Title: US-10-783-417-2

Perfect score: 3869  
Sequence: 1 NM00NDNNREYIIIDSHTSPY.....KLETTGKINTPFTNHTKTL 735

Scoring table:

BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 6049916 seqs, 41213615 residues

Total number of hits satisfying chosen parameters: 12099832

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-Processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Command line parameters:

-MODE=frame+ p2n.model -DEV=xlh  
-Q/cgn2\_1/USPTO.spool/US10783417/runat\_20012006\_095012\_23947/app\_query.fasta\_1.903  
-DB=Published Applications NA New -QMT=fastcap -SUFFIX=p2n.rnpbn -MINMATCH=0.1  
-LOOPEXT=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62  
-TRANS-human0.cdi -LIST=45 -DOCLIGN=200 -THR SCORE=pct -THR MAX=100  
-THR\_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pct -NORM=exc -HEAPSIZ=500 -MINLEN=0  
-MAXLEN=200000000 -USER=US10783417 @CGN\_1\_1121 @runat\_20012006\_095012\_23947  
-NCPU=6 -ICPU=3 -NO MMAP -NEG SCORES=0 -WAIT -DSPBLOCK=100 -DEV TIMEOUT=120  
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7  
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications NA New:\*

1: /cgn2\_6/ptodata/1/pubpna/US08\_NEW\_PUB.seq:\*  
2: /cgn2\_6/ptodata/1/pubpna/US06\_NEW\_PUB.seq:\*  
3: /cgn2\_6/ptodata/1/pubpna/US07\_NEW\_PUB.seq:\*  
4: /cgn2\_6/ptodata/1/pubpna/PCT\_NEW\_PUB.seq:\*  
5: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq:\*  
6: /cgn2\_6/ptodata/1/pubpna/US10\_NEW\_PUB.seq:\*  
7: /cgn2\_6/ptodata/1/pubpna/US11\_NEW\_PUB.seq1:\*  
8: /cgn2\_6/ptodata/1/pubpna/US11\_NEW\_PUB.seq2:\*  
9: /cgn2\_6/ptodata/1/pubpna/US11\_NEW\_PUB.seq3:\*  
10: /cgn2\_6/ptodata/1/pubpna/US11\_NEW\_PUB.seq4:\*  
11: /cgn2\_6/ptodata/1/pubpna/US60\_NEW\_PUB.seq:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	703	18.2	3633	8 US-11-058-727-3	Sequence 3, Appl1
2	703	18.2	3633	8 US-11-108-389-3	Sequence 3, Appl1
3	703	18.2	6613	8 US-11-058-727-18	Sequence 18, Appl1
4	689.5	17.8	6613	8 US-11-108-389-18	Sequence 18, Appl1
5	689.5	17.8	4188	10 US-11-091-643-5	Sequence 5, Appl1
6	678.5	17.5	2022	8 US-11-058-727-13	Sequence 13, Appl1
7	678.5	17.5	2022	8 US-11-108-389-13	Sequence 13, Appl1

8	675	17.4	3621	8 US-11-058-727-1	Sequence 1, Appl1
9	675	17.4	3621	8 US-11-108-389-1	Sequence 1, Appl1
10	675	17.4	4874	8 US-11-058-727-17	Sequence 17, Appl1
11	675	17.4	4874	8 US-11-108-389-17	Sequence 17, Appl1
12	669	17.3	4359	10 US-11-091-643-3	Sequence 3, Appl1
13	660	17.1	1959	8 US-11-192-801-3	Sequence 73, Appl1
14	657.5	17.0	2025	8 US-11-058-727-73	Sequence 73, Appl1
15	657.5	17.0	2025	8 US-11-058-727-79	Sequence 79, Appl1
16	657.5	17.0	2025	8 US-11-108-389-79	Sequence 73, Appl1
17	657.5	17.0	2025	8 US-11-108-389-79	Sequence 73, Appl1
18	657	17.0	2022	8 US-11-058-727-81	Sequence 81, Appl1
19	657	17.0	2022	8 US-11-108-389-81	Sequence 81, Appl1
20	656.5	17.0	2025	8 US-11-058-727-41	Sequence 41, Appl1
21	656.5	17.0	2025	8 US-11-058-727-47	Sequence 47, Appl1
22	656.5	17.0	2025	8 US-11-108-389-47	Sequence 41, Appl1
23	656.5	17.0	2025	8 US-11-108-389-47	Sequence 41, Appl1
24	656	17.0	1962	8 US-11-192-801-5	Sequence 5, Appl1
25	656	17.0	2022	8 US-11-058-727-49	Sequence 49, Appl1
26	656	17.0	2022	8 US-11-108-389-49	Sequence 49, Appl1
27	654.5	16.9	2022	8 US-11-058-727-69	Sequence 69, Appl1
28	654.5	16.9	2022	8 US-11-108-389-69	Sequence 69, Appl1
29	654	16.9	1959	8 US-11-192-801-1	Sequence 1, Appl1
30	653.5	16.9	2022	8 US-11-058-727-33	Sequence 33, Appl1
31	653.5	16.9	2022	8 US-11-058-727-67	Sequence 67, Appl1
32	653.5	16.9	2022	8 US-11-108-389-33	Sequence 33, Appl1
33	653.5	16.9	2022	8 US-11-108-389-67	Sequence 67, Appl1
34	653	16.9	2022	8 US-11-058-727-75	Sequence 75, Appl1
35	653	16.9	2022	8 US-11-108-389-75	Sequence 75, Appl1
36	652.5	16.9	2022	8 US-11-058-727-21	Sequence 21, Appl1
37	652.5	16.9	2022	8 US-11-058-727-63	Sequence 63, Appl1
38	652.5	16.9	2022	8 US-11-058-727-65	Sequence 65, Appl1
39	652.5	16.9	2022	8 US-11-108-389-21	Sequence 21, Appl1
40	652.5	16.9	2022	8 US-11-108-389-63	Sequence 63, Appl1
41	652.5	16.9	2022	8 US-11-108-389-65	Sequence 65, Appl1
42	652.5	16.9	2025	8 US-11-058-727-77	Sequence 77, Appl1
43	652.5	16.9	2025	8 US-11-108-389-77	Sequence 77, Appl1
44	652	16.9	2022	8 US-11-058-727-43	Sequence 43, Appl1
45	652	16.9	2022	8 US-11-108-389-43	Sequence 43, Appl1

#### ALIGNMENTS

RESULT 1  
US-11-058-727-3  
; Sequence 3, Application US/11058727  
; Publication No. US20050261483A1  
; GENERAL INFORMATION:  
; APPLICANT: Andre R. Abad  
; APPLICANT: Ronald D. Flanagan  
; APPLICANT: Rafael Herrmann  
; APPLICANT: Theodore W. Kahn  
; APPLICANT: Albert L. Lu  
; APPLICANT: Billy Fred McCutchen  
; APPLICANT: James K. Presnall  
; APPLICANT: James F. H. Wong  
; APPLICANT: Cao-Guo Yu  
; TITLE OF INVENTION: Genes Encoding Proteins With Pesticidal  
; TITLE OF INVENTION: Activity  
; FILE REFERENCE: 35718/287809  
; CURRENT APPLICATION NUMBER: US/11/058,727  
; CURRENT FILING DATE: 2005-02-15  
; PRIOR APPLICATION NUMBER: 60/391,786  
; PRIOR FILING DATE: 2002-06-26  
; PRIOR APPLICATION NUMBER: 60/460,787  
; PRIOR FILING DATE: 2003-04-04  
; PRIOR APPLICATION NUMBER: 10/606,320  
; PRIOR FILING DATE: 2003-06-25  
; NUMBER OF SEQ ID NOS: 134  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 3  
; LENGTH: 3633  
; TYPE: DNA  
; ORGANISM: Bacillus thuringiensis



Qy 599 -----ArgLeuGluIleThrCysGluThrProAsnSerThrGlnSerIlePheLeuArg 616  
Db 1705 CATATGCCTGTTAAATTCA-----GACATTACCAAGATATATGATGAG 17522  
Qy 617 LeuArgTyrAlaIleThrAsnGlyAlaGlyAsnThrLeuProAsnIleSerLeuThrIlePro 636  
Db 1753 ATTCGGATGCTCC-----GCTAATATACTGAAATTTATATTAATCTTCGAGAA 1806  
Qy 637 GlyValIleGlyIleProProGlnArgLeuAsnThrPheSerGlyThrAsnTyrAsn 656  
Db 1807 AACCTTAATCT---CACGCTCAAAAACCTATGATAGAGGTGAAGCTTTAAACATATAT 18633  
Qy 657 AsnLeuGlnTyrGlyAspPheGlyTyrPheGlnPheProSerThr-----ValIleThr 673  
Db 1864 AAATTTATATTATGGCACTTGGCCCCCTTATTAATTTACGACACCGAACTTTCAATATCT 192333  
Qy 674 LeuProLeuAsnArgAsnIleProPheIlePheAsnArgAlaAspValSerAsnSerIle 693  
Db 1924 CTAGGGGCT-----ATATTGAGCGGAAGACTTCTTGGAAATTGAA 19655  
Qy 694 LeuIleIleAspLysIleGluPheIleProIleThrIleThrSerMetHisGlnAsnArgGlu 713  
Db 1966 GCTTATATAGCCGAATCGAATTTATCCAGTAGATGAGACATAT-----GAAGCG 20166  
Qy 714 LysGlnLysLeuGluThrIleGluThrLysIleAsnThrPhePheThrAsn 730  
Db 2017 GACCAAGATTAGAAAGCAGCGAAGAAACGAGTAATGCTTGTATTCAGAT 2067

```

RESULT 2
US-11-108-389-3
; Sequence 3, Application US/11108389
; Publication No. US20050261186A1
; GENERAL INFORMATION:
; APPLICANT: Andre R. Abad
; APPLICANT: Ronald D. Flannagan
; APPLICANT: Rafael Herrmann
; APPLICANT: Theodore W. Kahn
; APPLICANT: Albert L. Lu
; APPLICANT: Billy Fred McCutchen
; APPLICANT: James K. Presnall
; APPLICANT: James F.H. Wong
; APPLICANT: Cao-Guo Yu
; TITLE OF INVENTION: Genes Encoding Proteins with Pesticidal
; TITLE OF INVENTION: Activity
; FILE REFERENCE: 35718/291049
; CURRENT APPLICATION NUMBER: US/11/108,389
; CURRENT FILING DATE: 2005-04-18
; PRIOR APPLICATION NUMBER: 60/391,786
; PRIOR FILING DATE: 2002-06-26
; PRIOR APPLICATION NUMBER: 60/460,787
; PRIOR FILING DATE: 2003-04-04
; PRIOR APPLICATION NUMBER: 10/606,320
; PRIOR FILING DATE: 2003-06-25
; NUMBER OF SEQ ID NOS: 134
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 3633
; TYPE: DNA
; ORGANISM: Bacillus thuringiensis
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(3633)
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (0)..(0)
; OTHER INFORMATION: Cg1218-2
US-11-108-389-3

Alignment Scores:
Pred. No.: 3,82e-68 Length: 3633
Score: 703.00 Matches: 219
Percent Similarity: 44.5% Conservative: 127

```

Best Local Similarity:	28.2%	Mismatches:	289
Query Match:	18.2%	Indels:	142
DB:	8	Gaps:	32

US-10-783-417-2 (1-735) X US-11-108-389-3 (1-3633)

[illegible]

Oy	343	AengluuThrArgGluIleThrGluIleAsnPhaArgLeuProGluLeu	362
Db	901	---CACTAACAGAGGAAGTATATCA-----GATCCACTGGCGCGGGT	942
Oy	363	ArgValGln-----ProAsnLeuAlaThrMetGluTy	373
Db	943	AACGGCTTCATAGTTGCTCGTATACAAAGACCTTCTTCGGAGTATAGATCA	1002
Oy	374	AsnLeuThrArgAlaSerPheLeuPheSerPheLeuGluGlnPheIlePheTy	393
Db	1003	TCGGTATATTCGACCAACC---CATGATATTGATATATATACGGGACTCACAGTGTATACA	1055
Oy	394	GluAsnThrAsnPheGlyAsn-----ArgLeuValGlyIleSerAsnArg	408
Db	1060	CAATCAAGAAGAGATTCTTCGCGCTGCATATATACACATGGCGCTGCATCAATAATAC	1113
Oy	409	AspAlaProThrTySerAsnThrIleThrGluThrLeuTyGlyGluArgThrGlySer	428
Db	1120	TATCATCGGATTTTATAGCATATATTTTAAACAGATGATAGCAATCAATCAATACTCA	1177
Oy	429	ProThrThrLeuThrIleArgProPheGluSerTy-----Lys	441
Db	1180	CACACACACTAGTACCTTT---GATTTACGAATTAAGATATTACACAGCTTATCAAAA	1233
Oy	442	ValSerIleValThrAspArgGlnSerProProValSerProIleGlnProHisPheIle	461
Db	1237	GATGGCGGTCCTCGTATATGTTTCTCGGTATACGATATATA-----TTTTTTGGA	1290
Oy	462	IleAsnGlnIleGluLeuTyIleAsnGlnGlySerSerAsnAsn-----ThrIleLys	478
Db	1291	ATGCCAGAGTCGAGGTTTTCATGGTAAACCAATTGAATTAACAGAAAGCGTTAAAG	1355
Oy	479	TyrSer-----AlaGlyGlySerLeuSerAsnTy	491
Db	1351	TATATCCGGTTCCAAAGATATTATATAGCGGG-----	1384
Oy	492	ThrPhePheGlnPheProArgIleLysAspCyAsnLeuValIleAspProGlyCysSer	511
Db	1384	-----ACAAAGAGTTCGGATTAAGATTAAGATTACCTCCAAACAACTTCA	1422
Oy	512	-----ProAsnPheAsnAsnTySerHisIleLeuSerHisPhe---SerLeuPheThr	528
Db	1423	GATCAACCAAAATTATAGTCAATATAGCCCAATATATATGTCATATACAGATATTTCCCGCG	1483
Oy	529	TyrSerTyValIleGlyLeuGlnLeuGlnIleLeuAspThrGlyValLeuGlyTyrThr	548
Db	1483	ACGGGTTCAACTACCGGATTA-----GTACCTGATATTTCTTGAGACA	1522
Oy	549	HisSerSerValAspArgTyTrAsnAlaIleSerAspLysIleIleThrMetIleProAla	568
Db	1525	CATCGAGAGCGCGAATCTTATAAATGACATTCAGATTAATAAATTATCAGATCCGGTC	1584
Oy	569	IleLysGlyAsnAsnLeu-----AspThrAsnSerLysValIleGlu	582
Db	1585	GTAAGAGTTTCGATTTGGCTCCCTCTATTAACAGAGAGGCCAAATATATACCGTTGATACG	1644
Oy	583	GlyProGlyPheIleThrGlyLysAsnLeuValTyLeuGlnSerGlnGly-----	598
Db	1645	GGTCCCTCGATTACAGGGGGGGGATTAATAAAGATTAAGAAAGAGATATTTATCA	1700
Oy	599	-----ArgLeuGluIleThrCysGluThrProAsnSerThrGlnSerTy	616
Db	1705	CATATGCGGTGTAAATTTCA-----GACATTACAAAGATATATGATGAGG	1755
Oy	617	LeuArgTyTrAlaThrAsnGlyAlaGlyAsnThrLeuProAsnIleSerLeuThrIlePro	636
Db	1753	ATTGGGTATGCTTCC-----GCTATATATATCTGAATTTATATATTAATCTCTGGAAGA	1800
Oy	637	GlyValIleIleGlyIleProProGlnArgLeuAsnAsnThrPheSerGlyTy	656
Db	1807	AACGTTAAATCT---CAGCTCAAAAAAATAATGAATAGAGTGAAGCTTTAAACATATAT	1863

```

QY      657 AenleuglnYrGlyAapPheNglYTrPhedInPheProSeTThr-----ValThr 673
           ||| ||||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
Db      1864 AAATTATTAATTTGTGCACCTTGCCCCCTTAAATTTACACACACCAGAACCTTTCAATTA 1923
           ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY      674 LeudProleuSnaYgaEnIlleProPheIllePheSnaIrgAlaApVaISerAenSerile 693
           ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      1924 CTACGGGGCT-----ATAATTGAACCGGAAGACTTTCTTGGAATTGAA 1965
           ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY      694 LeullellepPySllegluPheIlleProIleThreSerSeMethISglnAenArgGlu 713
           ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      1966 GCTTATATATAGCCGAATTCGAATTTATCCAGTAGATAGACATAT-----GAAGCG 2016
           ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY      714 lvsGlnIysleuGlnTrIlleglThrIylseIIeaSnThrPhePheThrAen 730
           ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      2017 GAACAAGAATTTHAGAACGACGCAAGCAAAAACAGTGAATGCCCTTGTTACGAAT 2067
           ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 3
US-11-058-727-18
; Sequence 18, Application US/11058727
; Publication No. US20050261483A1
GENERAL INFORMATION:
APPLICANT: Andre R. Abad
APPLICANT: Ronald D. Flannagan
APPLICANT: Rafael Herrmann
APPLICANT: Theodore W. Kahn
APPLICANT: Albert L. Lu
APPLICANT: Billy Fred McCutchen
APPLICANT: James K. Presnall
APPLICANT: James F.H. Wong
APPLICANT: Cao-Guo Yu
TITLE OF INVENTION: Genes Encoding Proteins With Pesticidal
FILE REFERENCE: 35718/287809
CURRENT APPLICATION NUMBER: US/11/058, 727
CURRENT FILING DATE: 2005-02-15
PRIOR APPLICATION NUMBER: 60/391, 786
PRIOR FILING DATE: 2002-06-26
PRIOR APPLICATION NUMBER: 60/460, 787
PRIOR FILING DATE: 2003-04-04
PRIOR APPLICATION NUMBER: 10/606, 320
PRIOR FILING DATE: 2003-06-25
NUMBER OF SEQ ID NOS: 134
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 18
LENGTH: 6613
TYPE: DNA
ORGANISM: Bacillus thuringiensis
FEATURE:
NAME/KEY: misc feature
LOCATION: (0)---(0)
OTHER INFORMATION: genomic Cry12Ib-2
US-11-058-727-18

Alignment Scores:
Pred. No.: 1.03e-67 Length: 6613
Score: 703.00 Matches: 219
Percent Similarity: 44.5% Conservative: 127
Best Local Similarity: 28.2% Mismatches: 289
Query Match: 18.2% Indels: 142
DB: 8 Gaps: 32

US-10-783-417-2 (1-735) x US-11-058-727-18 (1-6613)
QY      5 AsnAepAenAnGlnUTrGlnIlleIleAspSerIstHrSerProTyrrPheProAenArg 24
           ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      1263 AATATATCAAAATGATATAGAAATATAGATGCGACACTTCT-----ACTTCT 1310
           ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY      25 AsnSerAenAepSer----ArgTyrrProTyrrThrsAenSnProAenGlnProLeuGlnAen 43
           ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      1311 GTATTCGATGATGTTCTTACACGATACCTTTTCCGAAATGAGCCAACAATGCCCTTACAAAT 1370
           ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY      44 ThrsAntyrlvsgIutPrDleuAenMetCysGlnGlyAen--ThrglnTyrlvAaspAen 62
           ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

Db 1371 ATGAGTATTAAGATTATTTAAATAATGTCGGCGGAAATGCTAGTAATACCTCGGTCA 1430  
 QY phegluThrPheAlaSerAlaAspThrIleAlaIleValSerAlaGlyThrIleValSer 82  
 Db 1431 CCTGAAGTACTGTTAGCGGACAAAGT-----GCGAGCTAAGCGCGCAATTGATTAAGTA 1484  
 QY 83 GlyThrLeuLeuAlaGlyIleGlyIleuThrSerIleSerGlyProIleGlyIle 102  
 Db 1485 GGTAAATTACTATACGAGTTTAGG-----GTCCATTGTT 1520  
 QY 103 GlyAlaIleIleIleSerPheGlyThrLeuIleThrValPheThrProAlaGlyGluGln 122  
 Db 1521 GGGCCGATAGTAGTCTTAACTCACTTATGATATCTGTGGCTTCAGGGCA--- 1577  
 QY 123 AspIleThrValIleThrGlnIlePheIleuMetGlyGluIlePheValAspThrProLeu 142  
 Db 1578 ---AGAGTCATAGGAGATTTTATGACAAGTGAAGAACTCAATCAAAAATA 1634  
 QY 143 ThrGluSerIleuGlnLeuIleuGlnThrLeuGluGlyPheArgGlnIleLeuGln 162  
 Db 1635 GCGAATATGACAGAAATGAAGCGCTTGGAAATTGAAGAGATTAGATAATATTCAA 1694  
 QY 163 SerTyAsnThrAlaLeuAspAspArgIleuGlyIleuIleuArgLeuGlnAlaProGlyLeu 182  
 Db 1695 TTATATCTAAGCTCGCTTGAAGATGAAAGAA----- 1727  
 QY 183 ProProSerSerAlaLeuGlnGlnAlaAlaLeuThrLeuIleuValIleArgPheGluAsnVal 202  
 Db 1728 AATCCAAATGGTCA-----AGAGCTTAACGAGATGCGAATGGAATTTGAATTCG 1781  
 QY 203 HisAsnAspPheIleArgGluIleProGlyPheGlnLeuGlnThrTyIleThrLeuLeu 222  
 Db 1782 GATAGTTTATTTACGCAATACATGCCATCTTTGCGAGTACAATTTGAAGTACCATTC 1841  
 QY 223 LeuProIleTyAlaGlnAlaAlaAsnPheHisLeuAsnLeuLeuGlnGlnIleAlaGln 242  
 Db 1842 CTTCAGCTATATACACAGGACGCCAACCCTTCATTACTGTTATTAAGACCGTTCAAAT 1901  
 QY 243 LeuAlaAspGluIleTrpAsnAlaAspIleHisProSerGlnIleGluProAsnAlaGlyThr 262  
 Db 1902 TTTGGAGAAAGATGGCGATG-----TCTACAAACCACT 1934  
 QY 263 SerAspAspTyIleTyIleLeuLeuIleuGlnAsnIleProIleTyIleSerAsnTyIleVal 282  
 Db 1935 ATTAATACCTATATATATGCTCAATGAATGAACCTTACGACGAATATTTCTGATCACTGTGA 1994  
 QY 283 AsnThrTyIleArgThrGlyLeuIleuIleuAsnLeuArgAspGluProAsnMetIleTyIleSerIle 302  
 Db 1995 AAGCTGTAAGAACTGGTTTAGCAAAATTTAAAGGACGAGCGCTTAACATGGGTGCAC 2054  
 QY 303 PheAsnAspTyIleArgIleTyIleMetThrIleThrValIleuAspThrIleSerGlnPheSer 322  
 Db 2055 TATTAACCAATTCGCTGAGAAATGACATGACGCTTTAGATGTGTGCATATATCTCA 2114  
 QY 323 LeuTyIleAspIleuArgIleTyIleArgAspSerIleGlyIleGluValIleGlyIleIle 342  
 Db 2115 AATTATGACACAGCGATAC-----CCAATGGAACGAAGA----- 2153  
 QY 343 AsnGluLeuThrArgGluIleTyIleThrThrGlnIleAsnPheAspArgLeuProGlnLeu 362  
 Db 2154 ---CAACTAAGAGGAAGATATAC-----GATCACCCTGGCGCGCGGTA 2195  
 QY 363 ArgValGln-----ProAsnLeuAlaThrMetGluTyIle 373  
 Db 2196 AACGTCTCTTCAATGTGCTCTGATGACAAAGCACTCTTTCGAGATGATAGATAC 2255  
 QY 374 AsnLeuThrArgAlaSerPheIleuPheSerPheLeuGlnGlnIlePheTyIleThr 393  
 Db 2256 TCCGTATTTGACACACC---CATGTATTTGATTATTAACGGAAGCTCACAGTGTATAC 2312  
 QY 394 GluAsnThrAsnPheGlyAsn-----ArgLeuValGlyIleSerAsnArg 408  
 Db 2313 CAATCAAGAGCACTTCTCCGCTCGCTATATTAAGCATTTGGCTGTATCAAAATAGC 2372

QY 409 AspAlaProThrTyIleSerAsnThrIleThrGluThrLeuTyIleGlyAlaArgThrGlySer 428  
 Db 2373 TATCATCGGATTTTAGTATATATTTATTAAGATATAGAACTATGAAATCTA 2432  
 QY 429 ProThrThrTyIleThrIleArgProPheGluSerTyIle-----Ile 441  
 Db 2433 CACAGACTAGTACCTT---GATTTTACGATTTATGATTTTACAAAGCGTTATCAAA 2489  
 QY 442 ValSerIleValIleThrAspArgGlnSerProProValSerProIleGlnProHisPheIle 461  
 Db 2490 GATGGGCTGCTCTGATATGTTGTTTCTCGTTATACCTAATA-----TTTTTGA 2543  
 QY 462 IleAsnGlnIleGluLeuTyIleuAsnGlySerSerAsnAsn-----ThrIleuIle 478  
 Db 2544 ATGCAAGATGCAAGTTTTCATGTGAACCAATGAAATTAATCAAGAAAGCGTTAAG 2603  
 QY 479 TyIleSer-----AlaGlyIleSerIleSerAsnTyIleGlnAsnThr 491  
 Db 2604 TATTAATCCGCTTCCAAAGATATTTAGCGGG----- 2636  
 QY 492 ThrPhePheGlnPheProArgIleuIleuAspTyIleAsnLeuValIleAspProGlyCysSer 511  
 Db 2637 -----ACAAGAGATTCGGAATTTGAATTAACCTCAGAAACTTCA 2675  
 QY 512 -----ProAsnPheAsnAsnTyIleSerHisIleLeuSerHisPhe---SerLeuPheThr 528  
 Db 2676 GATCAACCAATTAATAGCATATACATAGATTTATGTCATATCAAGATATTCGCCGC 2735  
 QY 529 TyIleSerTyIleValIleGlyLeuGlnIleuGlnIleLeuAspThrGlyValIleuGlyThr 548  
 Db 2736 ACGGGTCAACACCGGATTA-----GTACCTGATTTTCTTGAC 2777  
 QY 549 HisSerSerValAspArgTyIleAsnAlaIleSerAspIleIleIleThrMetIleProAla 568  
 Db 2778 CATCGAGTGGCGATCTTAATATGCACTTCAATCAATTAATTAATCTCAGATTCGCGTC 2837  
 QY 569 IleIleGlyAsnAsnLeu-----AspThrAsnSerIleValIleGlu 582  
 Db 2838 GTAAAGTTTTCGATTTGGCTCCTCTATACAGAGGCGCAATTAATCCGTTGATCG 2897  
 QY 583 GlyProGlyHisIleThrGlyIleuAsnLeuValTyIleGlnSerGlnIle 598  
 Db 2898 GGTCTCGATTTTACAGGGGGGAGTATTAAGTAAGTAAGTAAGTAAGTAATTAATTAATCA 2957  
 QY 599 -----ArgLeuGlnIleThrCysGluThrProAsnSerThrGlnSerTyIlePheIleArg 616  
 Db 2958 CATATGCGGTAAATTAATTC-----GACATTAACAAAGATATAGTATGAGG 3005  
 QY 617 LeuAlaGlyTyIleThrAsnGlyAlaGlyAsnThrLeuProAsnIleSerLeuThrIlePro 636  
 Db 3006 ATTCGGTATGCTTCC-----GCTAATTAATCTGAATTTTATTAATATCTCTGGAAGA 3059  
 QY 637 GlyValIleGlyIleProGlnArgLeuAsnAsnThrPheSerGlyThrAsnTyIleAsn 656  
 Db 3060 AACGTTAAATCT---CACGCTAAAAAACTATGATAGGCGAAGCTTAACATATTAAT 3116  
 QY 657 AsnLeuGlnTyIleGlyAspPheGlyTyIlePheGlnPheProSerThr-----ValThr 673  
 Db 3117 AAATTAATTAATGCGACTTGGCCCCCTTAATTAATTAATTAACAAACGAACTTCAATTAAT 3176  
 QY 674 LeuProLeuAsnArgAsnIleProPheIlePheAsnArgAlaAspValSerAsnSerIle 693  
 Db 3177 CTAGGGGCT-----ATAATTTGAAGCGGAAGACTTCTTGGAAATTTGA 3218  
 QY 694 LeuIleIleAspIleIleGluPheIleProIleThrSerSerMetHisGlnAsnArgGlu 713  
 Db 3219 GCTTATATGACCGAATCAATTTATCCCAAGTATGACAAACATAT-----GAAGCG 3269  
 QY 714 TyIleGlnTyIleGlnIleThrIleGlnIleThrIleAsnThrPhePheThrAsn 730  
 Db 3270 GAACAAAGATTTAGAAAGCAAGAAAGCAAGAAAGCAAGTATGCTGTATTCGAAT 3320





```

QY 479 TyrSer-----AlaGlyGlySerLeuSerAsnTyrGlnAsnThr 491
DB 2604 TATTAATCCGGTTCCAAAGATATATATACGGGG----- 2636
QY 492 ThrPhePheGlnPheProArgLysLysAspCysAsnLeuValIleAspProGlyCysSer 511
DB 2637 -----ACAAAGATTCGGAATTGAAATTACCTCCAGAAACTTCA 2675
QY 512 -----ProAsnPheAsnAsnTyrSerHisIleLeuSerHisPhe---SerLeuPheThr 528
DB 2676 GATCAACCAAAATTAATGATCATATACCATAGATTATGTCATATACCAAGTATATCCCGG 2735
QY 529 TyrSerTyrValIleGlyLeuGlnLeuGlnIleLeuAspThrGlyValLeuGlyTyrThr 548
DB 2736 ACGGGTTCACACTACCGGATTA-----GTACCTGATATTTCTTGGAACA 2777
QY 549 HisSerSerValAspArgTyrAsnAlaIleSerAspLysIleIleThrMetIleProAla 568
DB 2778 CATCGGAGTGCAGATCTTATTAATGATGATTCATTCAGATTAATTAATCCGGTCC 2837
QY 569 IleLysGlyAsnAsnLeu-----AspThrAsnSerLysValIleGlu 582
DB 2838 GTAAAGGTTCTGATTGGCTCCCTCTATACAGAGGCGCAATATATACCGTTGTATCG 2897
QY 583 GlyProGlyHisThrGlyLysAsnLeuValTyrLeuGlnSerGlnGly----- 598
DB 2898 GGTCTCTGATTACAGGGGGGGGATATTAAGATTAAGAAATGAGTAATTAATATATCA 2957
QY 599 -----ArgLeuGlnIleThrCysGlnThrProAsnSerThrGlnSerTyrPheIleArg 616
DB 2958 CATATGCGGTTAATAATTCA-----GACATTAACAAGAAATATGATATGATGAGG 3005
QY 617 LeuArgTyrAlaThrAsnGlyAlaGlyAsnThrLeuProAsnIleSerLeuThrIlePro 636
DB 3006 ATTCGGTATGCTTCC-----GCTAATAATACTGAATTTATTAATAATCTCTGAAAGA 3059
QY 637 GlyValIleGlyIleProProGlnArgLeuAsnAsnThrPheSerGlyThrAsnTyrAsn 656
DB 3060 AACGTTAATCT---CACGCTCAAAAACATATGAAATAGAGGTAACCTTAAACATATAT 3116
QY 657 AsnLeuGlnTyrGlyAspPheGlyTyrPheGlnPheProSerThr-----ValThr 673
DB 3117 AAATTAATTAATGCGACTTGGCCCCCTATTAATTAATTAAGACAACGAACTTCTATTA 3176
QY 674 LeuProLeuAsnArgAsnIleProPheIlePheAsnArgAlaAspValSerAsnSerIle 693
DB 3177 CTAGGGGGCT-----ATATTGAAACCGAAGACTTCTTGGAATTGAA 3218
QY 694 LeuIleIleAspLysIleGlnPheIleProIleThrSerSerMetHisGlnAsnArgGlu 713
DB 3219 GCTTATATAGCCGAATCCAAATTTATCCAGTATGATGACACATAT-----GAAGCG 3269
QY 714 LysGlnLysLeuGlnIleGlnThrLysIleAsnThrPhePheThrAsn 730
DB 3270 GAACAAGATTGAAAGCAGCAAGAAAGACAGTGAATCCTGTTTACGAAT 3320

```

```

RESULT 5
US-11-091-643-5
; Sequence 5, Application US/11091643
; Publication No. US20050246789A1
; GENERAL INFORMATION:
; APPLICANT: TANAKA, Masao
; APPLICANT: YOKOYAMA, Tomoko
; APPLICANT: AOYAGI, Morichi
; APPLICANT: HASEGAWA, Makoto
; APPLICANT: EHARA, Gaku
; APPLICANT: KIMURA, Masaharu
; APPLICANT: NISHIHASHI, Hideji
; TITLE OF INVENTION: Polypeptide having larvae growth inhibiting or
; TITLE OF INVENTION: Insecticidal effect on scabiaeidae insects and
; FILE REFERENCE: OP1335

```

```

; CURRENT APPLICATION NUMBER: US/11/091,643
; CURRENT FILING DATE: 2005-03-29
; PRIOR APPLICATION NUMBER: JP 2001-115754
; PRIOR FILING DATE: 2001-04-13
; PRIOR APPLICATION NUMBER: JP 2001-203463
; PRIOR FILING DATE: 2001-07-04
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
; LENGTH: 4188
; TYPE: DNA
; ORGANISM: Bacillus popilliae
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(4158)
US-11-091-643-5

Alignment Scores:
Pred. No.: 1,65e-66 Length: 4188
Score: 689.50 Matches: 239
Percent Similarity: 45.1% Conservative: 113
Best Local Similarity: 30.6% Mismatches: 286
Query Match: 17.8% Indels: 143
DB: Gaps: 38

US-10-783-417-2 (1-735) x US-11-091-643-5 (1-4188)

QY 1 MetAsnGln-----AsnAsnAspAsn-----AsnGluTyr 10
DB 55 ATGAAATCATATCATATACCAACCAATCAAAAGTACAAACCAAGTGAATGAAATGAG 114
QY 11 GlnIleIle-----AspSerHisThrSerProTyrPheProAsnArgAsnSerAsn 28
DB 115 CAATTAATCAACCTTCAAGTAAAGCTTACTTTACAGTCCCAAC----- 159
QY 29 SerArgTyrProTyrThrAsnAsnProAsnGlnProLeuGlnAsnThrAsnTyrLysGlu 48
DB 160 ---AGTATCCGTAATGCTACGATCCCAATGCTATCATACGAGAGGTATGAAATTAAT 216
QY 49 TrpLeuAsnMetCysGlnGlyAsnThrGlnTyrGlyAspAsnPheGlnThrPheAlaSer 68
DB 217 TGGTGGATATGATGTGTAGGT-----GTAGGTGACGAT-----ACA 252
QY 69 AlaAspThrIleAlaValSerAlaGlyThrIleValSerGlyThrLeuLeuAlaGly 88
DB 253 CGAAGTCCCGAAGCTCGGGTATCTGCACAAAGTCTATTTCACATCTCTT----- 303
QY 89 IleGlyGlyLeuThrSer---IleSerGlyProIleGlyIle-----IleGlyAlaIle 105
DB 304 -----GGATTAACAAGACATCATATGCGCTCATAGATATTCGCGTTGAGCGCAAGCC 357
QY 106 IleIleSerPheGlyThrIleIleThrValPheTrpProAlaGlyGlnGlnAspLysThr 125
DB 358 ATTTGAATTTTGGCGCACTATGATTTGTGTGGCTGCGACGCGCTGATCCA----- 411
QY 126 ValTyrThrGlnPheIleLysMetGlyGlnIlePheValAspThrProLeuThrGlnSer 145
DB 412 ---TGGGTATATTTATGATCATCATGTAGAAAGACTCATTAATTTCAAAATTAACAGACT 468
QY 146 IleLysGlnLeuLysLeuGlnThrLeuGlnGlyPheArgGlnIleLeuGlnSerTyrAsn 165
DB 469 GTAAATAATGAGCAATTAACAAGATTTAGACGTTTATAGCAATGCTCTAGCGCTATATCA 528
QY 166 ThrAlaLeuAspAspTyrArgLysLeuLysArgLeuGlnAlaProGlyLeuProProSer 185
DB 529 AAGGCTTTCAGAGATGCGCA-----CAACATCT----- 558
QY 186 SerAlaLeuGlnGlnAlaAlaLeuThrLeuLysIleArgPheGlnAsnValHisAsnAsp 205
DB 559 ---ACTTTGAATCAAGCTCGACTACGTGTAACAGATGATTTTCTTAATGTAATAATTT 615
QY 206 PheIleArgGlnIleProGlyPheGlnLeuGlnGlnThrTyrLysThrLeuLeuProIle 225

```

Db 616 TTTGAGGACCTTATGCTTCATTCAGAGTACCAAGCTTATGAAGTACCATTTATTAACGCTA 675  
 Qy 226 TTTAAGTAAATTAAsnPhenileuAsnleuEngInglYAlaGluLeuAlaasp 245  
 Db 676 TACGATACCGGTCAAACTCCATTTATTTATTAAGAGATAGCCGATTTTCGGTTTG 735  
 Qy 246 GUTTPAsnAlaaspIleHisProSerGInIleGluProAsnAlaGlyThrSerAsp 265  
 Db 736 GATTGGGGATTA-----AGTCAAACTCATGTTAAC-----GAT 768  
 Qy 266 TTTTTLyLeuLeuLeuYsgIuAsnIleProLyThrSerAsnThrCysAlaAsnThrTy 285  
 Db 769 AATTATATCTCCAAATTAAGCGCTTCGACATTTATGCAATATGTAACAATCTGGTAT 828  
 Qy 286 ArgThrGlyLeuLyAsnleuArgAspGluProAsnMetLysTyrSerIlePheAsnAsp 305  
 Db 829 CGGACGGGTTTACAAAGATTGCAAGGACCAATGTCAGACAGTTGGGTCAATTATATCA 888  
 Qy 306 TTTATGAGTyrMetThrIleThrValIleuAspThrIleSerGlnPheSerLeuTyrAsp 325  
 Db 889 TTTAGAGAGAAATGACACTAAGTATAGATGTTTGTCATTTATTTCAAGTTATGAT 948  
 Qy 326 IleLyArgTyrArgAspSerIleGlyGlyIleGluValIleGlyTyrIleLyAsnGluLeu 345  
 Db 949 TATCGAGTTAC-----CCAAATGAGCTTAAGGGA-----GAGCTT 984  
 Qy 346 ThrArgGluIleTyrThrThrGluIleAsnPheAspArgLeuProGlnleuArgValGln 365  
 Db 985 ACGAGAGAAATTTATACGATCCAGTACAGAGAGCTCTTT-----TGGGTAAATGAGACA 1038  
 Qy 366 ProAsnleuAlaThrMetGluTyrAsnleuThrArgAlaSerPheLyLeuPheSerPhe 385  
 Db 1039 CCAAACTTCGACATAGAAATAACAGTAGTAGGCAACCAACACCC-----TTTACTTGG 1095  
 Qy 386 LeuGluGlnPheIlePheTyrThrGluAsnThrAsnPheGlyAsnArgLeuValGlyIle 405  
 Db 1096 CTAGTTACTTTAACAGTTAATACAGTCAAGTACAGTCTGGCGAT-----GGAAT 1146  
 Qy 406 SerAsnArgAspAlaProThrTyrSerAsnThrIleThrGluThrLeuTyrGlyGluArg 425  
 Db 1147 TCTAATCTATTTATGGAATCACATAGTCAACCGTAGAGAAACCGAGAGCTCAGTCT 1206  
 Qy 426 ThrGlySerProThr-----ThrLyThrIleArgProPheGluSerTyrLyVal 442  
 Db 1207 ATTCAAGAGTCTTACCTGAGAGTACTGTCACAAT-----TATGCG-- 1248  
 Qy 443 SerIleValThrAspArgGln-----SerProProValSerProIleGlnProHisPhe 460  
 Db 1249 -----ACGATATATTACTTTTAAATCA-----TTT 1275  
 Qy 461 IleIleAsnGlnIleGluLeu-----TyrLeuAsnGlySerSerAsnAsnThr 476  
 Db 1276 TTATTATAGGATATTATTAACCATTAATACAGGTATAGTTTCTTATCTGCTAATTTGTT 1335  
 Qy 477 LeuLyTyrSerAlaGlyGlySerLeuSerAsn-----TyrGlnAsnThr 491  
 Db 1336 GGAATCTATTCAGTATTTACAGCACTCGTTCTATTAAGCTTCTGTATAGAACCA 1395  
 Qy 492 ThrPhePheGlnPheProArgLySlyAspCysAsnLeuValIleAspProGlyCysSer 511  
 Db 1396 AAGATTTTCCAGCTTACATATCAATTTCTGTAATTA-----CTGAGATTAAC 1446  
 Qy 512 -----ProAsnPheAsnAsnTyrSerHisIleLeuSerHisPheSerLeuPheThr 528  
 Db 1447 TCGGATAGGCCCAACTGCTCCGACTATATCTATAGCTATCTGATATCTCAGGTTTGA 1506  
 Qy 529 TTTSerTyrValIleGlyLeuGlnleuGlnIleLeuAspThrGly-----Val 544  
 Db 1507 -----ACTGATGGGAGAGACGGTTCTAGTT 1533  
 Qy 545 LeuGlyTTPThrHisSerSerValaAspArgTyrAsnAlaIleSer-----AspLyIleIle 563  
 Db 1534 TATGGTGGACATCTTCAACTGCTACTGAGAAATATTAACGTTAGACAGACAGAAATA 1593

Qy 564 ThrMetIleProAlaIleLyGlyAsnAsnleuAspThrAsnSerLyLeuValIleGluGly 583  
 Db 1594 GTACAACTTCCAGCTGTTAAGGAAACAAGTCTCAAC--AATTGCCAAGTAGTTAGAGGA 1650  
 Qy 584 ProGlyHisThrGlyGlyAsnleuValTyrleuGlnSerGlnGlyArgLeuGluIleThr 603  
 Db 1651 ACTGATTTTACAGGAGGAGCTGCTTGAAGCTTAATATGTAACATTTTCTTACT 1710  
 Qy 604 CysGluThrProAsnSerThrGlnSerTyrPheIleArgLeuArgTyrAlaThrAsnGly 623  
 Db 1711 CTGGTTTC--AGTCACTTACCTTACCGCCCTCCGATTCGTTATGCTCCGACGA 1767  
 Qy 624 AlaGlyAsnThrLeuProAsnIleSerLeuThrIleProGlyValIleGlyTyrProPro 643  
 Db 1768 GCGCGATCA-----GGTTTCTCTGTTATATGGAATCAATATGAGAAATTTCCA 1818  
 Qy 644 GlnArgLeuAsnAsnThrPheSerGlyThrAsnTyrAsn-----AsnleuGlnTyr 660  
 Db 1819 --ACCACACAGATATCCCTTCTCCCAATGTACTCACTGCCCAAAATGTACCATAC 1875  
 Qy 661 GlyAspPheGlyTyrPheGlnPheProSerThrValThrLeu----- 674  
 Db 1876 GAGCTTTTAAAGATTGTAGATTATTAATCTTACTGTACTTATTAAGAAATATCTTCTGCT 1935  
 Qy 675 -----ProLeuAsnArgAsnIleProPheIlePheAsnArgAlaAspValSerAsnSer 692  
 Db 1936 TCAACAACTTTTCGACTGATTTCCGTTTCATTTGTCATTAGAAATACTCGCAAT-- 1992  
 Qy 693 IleleuIleIleAspLyIleGluPheIleProIleThrSerSerMetHisGlnAsnArg 712  
 Db 1993 --ATATTAATTTGACCAATTAATTTGTCCCATTTAGAGGTTCTCTTTCGAGTACGA 2049  
 Qy 713 GluLyGlnLyLeuGluThrIleGlnThrLyIleAsnThrPhePheThrAsnHisThr 732  
 Db 2050 ACCAAACGACGCTTGAAGAAAGCAAGAAAGCGGTAACCATTTGTTACAGATGATCG 2109  
 Qy 733 Lys 733  
 Db 2110 AAA 2112

RESULT 6  
 US-11-058-727-13  
 ; Sequence 13, Application US/11058727  
 ; Publication No. US20050261483A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Andre R. Abad  
 ; APPLICANT: Ronald D. Flannagan  
 ; APPLICANT: Rafael Herrmann  
 ; APPLICANT: Theodore W. Kahn  
 ; APPLICANT: Albert L. Lu  
 ; APPLICANT: Billy Fred McCutchen  
 ; APPLICANT: James K. Prensall  
 ; APPLICANT: James F.H. Wong  
 ; APPLICANT: Cao-Guo Yu  
 ; TITLE OF INVENTION: Genes Encoding Proteins with Pesticidal  
 ; FILE REFERENCE: 35718/287809  
 ; CURRENT APPLICATION NUMBER: US/11/058,727  
 ; PRIOR FILING DATE: 2005-02-15  
 ; PRIOR APPLICATION NUMBER: 60/391,786  
 ; PRIOR FILING DATE: 2002-06-26  
 ; PRIOR APPLICATION NUMBER: 60/460,787  
 ; PRIOR FILING DATE: 2003-04-04  
 ; PRIOR APPLICATION NUMBER: 10/606,320  
 ; PRIOR FILING DATE: 2003-06-25  
 ; NUMBER OF SEQ ID NOS: 134  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 13  
 ; LENGTH: 2022  
 ; TYPE: DNA  
 ; ORGANISM: Bacillus thuringiensis (truncated)  
 ; FEATURE:



Oy	599	-----AGLeagLlIeTh-CyeGluThrProAsnSerThrInseryrPheIIeArg	616
Dd	1705	CATATGCCTTTAAATTTC-----GACATTACAAGAATATAGTATGAG	1752
Oy	617	LeuArGlyAlaIarMaenGlyAlaGlyAsnThrIeuProasnlleSeryleThrIllePro	636
Dd	1753	ATTGGTAATGCTCC-----GCTAATATATCTGAATTTATAAACCTCTCGAAGAA	1806
Oy	637	GlyValIIegLyIleProProGIaThArgIeuAsnAnthrPheserGlyThrAsnTyAsn	656
Dd	1807	AACGTTAATCT---CAGCTCAAAAACACTGATTAGAGTGAAAGCTTTAACTATAT	1863
Oy	657	AsnIeuGlnTyrgLyAspPheGlyTyrrheGlnPheProserThr-----ValThr	673
Dd	1864	AAATTTAATTTAGCATCTTGCCCCCTTAAATTTAGACAAACCACTTTCACTACT	1923
Oy	674	IeuProIeuAsnArgsnlleProPheIlePheAsnArgLaaspValiserAsnserile	693
Dd	1924	CTAGGGCGCT-----ATATTGACGCCGAAGACTTCTTGGAATTGAA	1965
Oy	694	LeuIlelleaspLysllegInPheIIeProIle	704
Dd	1966	GCTTATATAGACCGAATCGAATTTATCCACAGTA	1998

QY	363	ArgValGln-----	-----ProAsnLeuAlaThrMetGluIyr	373
Db	943	AACGGCTCTCAATGTTGCTTCGTATGACAAAGCAACCTTCTTCGGAGTGAATGAATCA	1002	
QY	374	AsnLeuThrArgAlaSerPheLysLeuPheSerPheLeuGlnGlnPheIlePheYrThr	393	
Db	1003	TCCGTTATTGCAGCACCC-----CATGATATTGATTATATACGAGACTCACAGTGTATAC	1055	
QY	394	GlnLeuThrAsnPheGlyAsn-----	-----ArgLeuValGlyIleSerAsnArg	408
Db	1060	CAATCAAGAAGACATTTCTTCGCTCGCTATATATACAGATTCGGCTGATCAATAATAC	1119	
QY	409	AspAlaProThrYrSerAsnThrIleThrGluThrLeuYrGlyGlnArgThrGlySer	428	
Db	1120	TATCATCGGAATTTTATGATATATATTTATTAACAGATGTATAGCAATATCAAAATCTA	1179	
QY	429	ProThrThrLysThrIleArgProPheGluSerYr-----	-----Lys	441
Db	1180	CACAGACACTAGTACCTTT-----GATTTTACGAATTTGATATTATACAAACGTTATCAAA	1233	
QY	442	ValSerIleValThrAspGlnSerProValSerProIleGlnProHisPheIle	461	
Db	1237	GATCGGGAGCTCCTTGATATGTTTCTTCGTATATAGTATATA-----TTTTTGG	1299	
QY	462	IleAsnGlnIleGluLeuYrLeuAsnGlySerSerAsnAsn-----	-----ThrLeuLys	478
Db	1291	ATGCAGAGATCGAGTTTTCATGCTGTAAACCAATGAAATTAATACAGAAAGCGTTAAAG	1356	
QY	479	TySer-----	-----AlaGlyGlySerLeuSerAsnYrGlnAsnThr	491
Db	1351	TATATATCCGGTTTCCAAAGATATATATAGCGGG-----	-----	1388
QY	492	ThrPhePheGlnPheProArgLysLysAspCysAsnLeuValIleAspProGlyCysSer	511	
Db	1384	-----ACAAAGAGTTTCGGAATTAAGATTAAGATTAATCTCCGAACCTCA	1422	
QY	512	-----ProAsnPheAsnAsnYrSerHisIleLeuSerHisPhe--	-----SerLeuPheThr	528
Db	1423	GATCAACCAAAATTAATGACATATATGCCATATGATATATGATATACAAAGTATATCCCGC	1482	
QY	529	TyrSerYrValIleGlyLeuGlnLeuGlnIleLeuAspThrGlyValLeuGlyYrPheThr	548	
Db	1483	ACGGGTTCACTACCGGATTA-----	-----GTACCTGATTTTCTTGACA	1522
QY	549	HisSerSerValAspArgYrAsnAlaIleSerAspLysIleIleThrMetIleProAla	568	
Db	1525	CATCGGAGTCCCGGATCTTATATATGCAATCTTTCAGTAAATTAATCTCAGATTCGGGTC	1586	
QY	569	IleLysGlyAsnAsnLeu-----	-----AspThrAsnSerLysValIleGlu	582
Db	1585	GTAAGGTTTCTGATTTGGCTCCCTCTATTAACAGAGAGGCGCAAAATTAATACGTTGTATCG	1644	
QY	583	GlyProGlyHisThrGlyGlyAsnLeuValYrLeuGlnSerGlnGly	598	
Db	1645	GCTCTCGAATTTACAGGGGGGGGGAATAAAGTAATTAAGAAATGAGTAATTATATCA	1704	
QY	599	-----ArgLeuGlnIleThrCysGluThrProAsnSerThrGlnSerYrPheIleArg	616	
Db	1705	CATATCGCGTTAAATTTCA-----	-----GACATTAACAAAGATATATGATGAGC	1755
QY	617	LeuArgYrValaThrAsnGlyValaGlyAsnThrLeuProAsnIleSerLeuThrIlePro	636	
Db	1753	ATTCCGATGCTTCC-----GCTATATATATCTGAATTTTATATTAATCTCTGAGAA	1806	
QY	637	GlyValIleGlyYrIleProProGlnArgLeuAsnAsnThrPheSerGlyYrThrAsnYrAsn	656	
Db	1807	AACGTTAAATCT--CAAGCTCAAAAACATATGAATTAAGGTGAAGCTTTAACATATATAT	1865	
QY	657	AsnLeuGlnYrGlyAspPheGlyYrPheGlnPheProSerThr-----	-----ValThr	673

```

Db      1924 CTAGGGGCT-----ATATTTTGAAGCGAAGACTTTCTTGGAATTGAA 1965
QY      694 LeuileleaplysliegIuPheileProile 704
      |||:::|||||:::
Db      1966 GCTTATATAGACCGAATGCAATTATTCACAGTA 1998

RESULT 8
US-11-058-727-1
/ Sequence 1, Application US/11058727
/ Publication No. US20050261483A1
/ GENERAL INFORMATION:
/ APPLICANT: Andre R. Abad
/ APPLICANT: Ronald D. Flannagan
/ APPLICANT: Rafael Herrmann
/ APPLICANT: Theodore W. Kahn
/ APPLICANT: Albert L. Lu
/ APPLICANT: Billy Fred McCutchen
/ APPLICANT: James K. Preenail
/ APPLICANT: James F.H. Wong
/ APPLICANT: Cao-Guo Yu
/ TITLE OF INVENTION: Genes Encoding Proteins with Pesticidal
/ TITLE OF INVENTION: Activity
/ FILE REFERENCE: 35718/287809
/ CURRENT APPLICATION NUMBER: US/11/058, 727
/ CURRENT FILING DATE: 2005-02-15
/ PRIOR APPLICATION NUMBER: 60/391, 786
/ PRIOR FILING DATE: 2002-06-26
/ PRIOR APPLICATION NUMBER: 60/460, 787
/ PRIOR FILING DATE: 2003-04-04
/ PRIOR APPLICATION NUMBER: 10/606, 320
/ PRIOR FILING DATE: 2003-06-25
/ NUMBER OF SEQ ID NOS: 134
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 1
/ LENGTH: 3621
/ TYPE: DNA
/ ORGANISM: Bacillus thuringiensis
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: (1)...(3621)
/ FEATURE:
/ NAME/KEY: misc feature
/ LOCATION: (0)...(0)
/ OTHER INFORMATION: Cpy1218-1
US-11-058-727-1

Alignment Scores:
Pred. No.: 5,78e-65 Length: 3621
Score: 675.00 Matches: 216
Percent Similarity: 44.5% Conservative: 125
Best Local Similarity: 28.2% Mismatches: 300
Query Match: 17.4% Indels: 126
DB: 8 Gaps: 32

US-10-783-417-2 (1-735) x US-11-058-727-1 (1-3621)
QY      5 AenbapbanaangIuTYrgIuileleapSerHisThrSerProTYrPheProAsnArg 24
      |||:::|||||:::
Db      10 AATATCAAAATGAAATATGAAATATGATGCGACACCTTCT-----ACTTCT 57
QY      25 AenSerAenApsSer---ArgTYrProTYrThrAsnAsnProAsnAngIuProleuGlnAsn 43
      |||:::|||||:::
Db      58 GTATCCAAATGATCTTAACAGATACCCCTTTTCGATGAGCCACAAATGCGCTACAAAT 117
QY      44 ThrAsnTYrLysgluTYrLeuAsnMetCysglngIyAsn---ThrgIuTYrgIyAspAsn 62
      ::::|:::|||||:::
Db      118 ATGATATTATAAAGATTATTATAAATATGTCGCGGAATGCTAGTGAATACCTGATTA 177
QY      63 PhegluThrPheAlaSerAlaAspThrIleAlaIaIaValSerIaagIyThrIleValSer 82
      |||:::|||||:::
Db      178 CCTGAAGTACTTGTGAGGACAAGAT-----GCACCTAAGGCCGAATGATGATATGTA 231

```



```

? APPLICANT: Ronald D. Flannagan
? APPLICANT: Rafael Herrmann
? APPLICANT: Theodor W. Kahn
? APPLICANT: Albert L. Lu
? APPLICANT: Billy Fred McCutchen
? APPLICANT: James K. Presnail
? APPLICANT: James F.H. Wong
? APPLICANT: Cao-Guo Yu
? TITLE OF INVENTION: Genes Encoding Proteins With Pesticidal
? TITLE OF INVENTION: Activity
? FILE REFERENCE: 35718/291049
? CURRENT APPLICATION NUMBER: US/11/108,389
? CURRENT FILING DATE: 2005-04-18
? PRIOR APPLICATION NUMBER: 60/391,786
? PRIOR FILING DATE: 2002-06-26
? PRIOR APPLICATION NUMBER: 60/460,787
? PRIOR FILING DATE: 2003-04-04
? PRIOR APPLICATION NUMBER: 10/606,320
? PRIOR FILING DATE: 2003-06-25
? NUMBER OF SEQ ID NOS: 134
? SOFTWARE: FastSeq for Windows Version 4.0
? SEQ ID NO 1
? LENGTH: 3621
? TYPE: DNA
? ORGANISM: Bacillus thuringiensis
? FEATURE:
? NAME/KEY: CDS
? LOCATION: (1)...(3621)
? FEATURE:
? NAME/KEY: misc_feature
? LOCATION: (0)...(0)
? OTHER INFORMATION: Cry12Ia8-1
? US-11-108-389-1

```

Pred. No.:	5,786-65	Length:	3,631
Score:	675.00	Matches:	216
Percent Similarity:	44.5%	Conservative:	125
Best Local Similarity:	28.2%	Mismatches:	300
Query Match:	17.4%	Indels:	126
DB:	8	Gaps:	32

US-10-783-417-2 (1-735) x US-11-108-389-1 (1-3621)

QY	5	AsnaPaasnaSnglYrGluIleIleAspSerHisThrSerProTYrPheProAsnArg	24
Db	10	AATAATCAAAATGAATATGAAATTATATGATGCACACCTCT	57
		-----ACTTCT	
QY	25	AsnSerHisAspSer---ArgTYrProTYrThrAsnAsnProAsnGlnProLeuGlnAsn	43
Db	58	GTATCCAAATGATTCCTAACAGATAACCTTTTGCAATGAGCAACAAATGCGCTACAAAT	117
QY	44	ThraenTYrLeGluTYrLeuAsnMetCysGlnGlyAsn---ThrglnTYrGlyAspAsn	62
Db	118	ATGATATTATTAAGATTATTATTTAAAAATGTCGCGGAAATCTGTGTAATCCCTGATTCA	177
QY	63	PheGluThrPheAlaSerAlaAspThrIleAlaIleValSerAlaGlyThrIleValSer	82
Db	178	CCTGAAGTACTTGTTGCGGACAAAGAT-----GCAGCTAAGGCCCAATTGATATGATA	231
		-----	
QY	83	GlyThrIleuLeuAlaGlyIleGlyGlyLeuThrSerIleSerGlyProIleGlyIleIle	102
Db	232	GGAATATTACATACAGGTTAGGG-----GTCCCATTTGTT	267
		-----	
QY	103	GlyAlaIleIleIleSerPheGlyThrIleuIleThrValPheThrProAlaGlyGln	122
Db	268	GGGCGGATAGTAGCTTATATACTCAACTATGATATTCGTGGCCTTCAGGGAA---	324
QY	123	AspGluThrValTYrPheGlnPheIleIleYrMetGlyGluIlePheValAspThrProLeu	142
Db	325	--AAGAGTCATATGGAAATTTTATATGGAACAAAGTAAAGAACTCAATTATCAAAAATA	381
QY	143	ThrgIleSerIleGlyGlnIleuYrSleuGlnThrLeuGlnGlyPheArgGlnIleLeuGln	162

Dd	382	GCAGAAATATGCAGAAATTAAGCCCTTTCCGAAATTAGAGATTAGGTAATATTACCA	441
Oy	163	SeTyrAsnThrAlaLeuAspAspTrpArgLysLeuYsArgLeuGlnAlaProGlyLeu	182
Dd	442	TTAATATCAATCGCGCTTGAAGATGGGAGAA-----	474
Oy	183	ProProSerSerAlaLeuGlnGlnAlaAlaLeuThrLeuYsIleArgPheGluAsnVal	202
Dd	475	AATCCAATAGGTCTCA-----AGAGCCTTACGAGATGTGCGAATAATTCATTGGAATCTCG	528
Oy	203	HisAsnAspPheIleArgLuuIleProGlyPheGluLeuGluThrTyrLysThrLeu	222
Dd	529	GATGGTTATTATTACGCAATATATATGCCATTATTAGAGTGCACAAATTTTGAAGTACATTTC	588
Oy	223	LeuProIleTyrAlaGlnAlaAlaAspPheHisLeuAsnLeuLeuGlnGlnAlaGlu	242
Dd	589	CTTACTGTATATGCGAATGCGACGCCACACTTCATTATCTGTTATTAAAGACCGCTCAATT	648
Oy	243	LeuAlaAspGluTrpAsnAlaAspIleHisProSerGluIleGluProAsnAlaGlyThr	262
Dd	649	TTTGAGAGAGATGGGATGG-----TCACAACATCTT	681
Oy	263	SerAspAspTyrTyrLysLeuLeuLeuGluAsnIleProLysTyrSerAsnTyrCysAla	282
Dd	682	ATTATATACATTTTATGATCGTCAATGAAATCTTACTGCAGAAATTTCTGATCACTGTGTA	741
Oy	283	AsnThrTyrArgThrGlyLeuLysAsnLeuArgAspGluProAsnMetLysTrpSerIle	302
Dd	742	AAGTGATATGAACCTGGTTTACGCAAAATTAATAAGCGACGACCGCTTAACATGGGCTTGAC	801
Oy	303	PheAsnAspTyrArgArgTyrMetThrIleThrValLeuAspThrIleSerGlnPheSer	322
Dd	802	TATATCACAATTCGGTAGAAGAAATGACACTGGCGGTTTAAATGTTGTTGATTTATCCCA	861
Oy	323	LeuTyrAspIleLysArgTyrArgAspSerIleGlyGlyIleGluValIleGlyIleLys	342
Dd	862	AATTATATGACACCCACGATAC-----CCAAATGGAACGAAAGCA-----	900
Oy	343	AsnGluLeuThrArgGluIleTyrThrGluIleAsnPheAspArgLeuProGluLeu	362
Dd	901	---CACTACACAAAGGAGATATATACA-----GATCCACTGGCGCGGTA	942
Oy	363	ArgValGln-----ProAsnLeuAlaThrMetGluTyr	373
Dd	943	AACGTGTCTTCAATTGGTCTCGGTATAGACAAACACCTTCTTCGAGTATGAAATCA	1002
Oy	374	AsnLeuThrArgAlaSerPheLysLeuPheSerPheLeuGluGluInPheIlePheTyrThr	393
Dd	1003	TCCGTATTTCGACACACC---CATGTATTGATATATTAACGGGACATCACAGTGTATATCA	1059
Oy	394	GluAsnThrAsnPheGlyAsn---ArgLeuVal-----GlyIleSer	406
Dd	1060	CAATTCAGAAACATTTCTTCGCTCGGTATATTAAGACATGGCGTGCATCAAAATATAGC	1119
Oy	407	AsnArgAspAlaProThrTyrSerAsnThrIleThrGluThrLeuTyrGlyGluArgThr	426
Dd	1120	TACCATCGTGCAGTAGGGGTATAT-----CTTCAACAAATGTATGAACTATATCA	1173
Oy	427	GlySerProThrThrLysThrIleArgProPheGluSerTyrLysValSerIleValThr	446
Dd	1174	AATCTACACAGCACTAGTACCTTT---GATTTTACGATATATGATATTACAGACCTTCA	1230
Oy	447	AspArgGlnSerProProValSerProIleGlnPro-----HisPheIle	461
Dd	1231	TCAAGAGATGACACTCTCTTGATATTTGATATCCCTGGTATATACGTATATATTTTGGCA	1290
Oy	462	IleAsnGlnIleGluLeuTyrLeuAsnGlySerSerAsnAsn-----ThrLeuLys	478
Dd	1291	ATGCCAGAAAGTCAGGTTTTCATGTGTAACCAATTTGAATATATACAGAAAGCCTTAAG	1350
Oy	479	TySerAlaGlyGlySerLeuSerAsnTyrGlnAsnThrThrPhePheGlnPheProArg	498





```

Db      1319 CTTACTGATATGCAATGAGGACCACTTACTGTTATTAAGACGGCTCAATT 1378
Qy      243 LeuAlaAspGluThrPheAlaIleHisProSerGlnIleGluProAlaIleGlyThr 262
Db      1379 TTTGAGAGAAATGGGGATGG-----TCACCACTACT 1411
Qy      263 SerAspAspTyrTyrIleuLeuLeuGluAlaIleProIleTyrSerAlaIleCysAla 282
Db      1412 ATTAATTAATCTATTATGATGCTCAATGAATGAACTTACGAGAAATATCTGATCTGTA 1471
Qy      283 AenThrTyrArgThrGlyLeuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleu 302
Db      1472 AAGGTGATGAAGAACTGTTTACGAAATTAAGACGAGCGCTAAACATGGGTGAC 1531
Qy      303 PheAsnAspTyrArgArgTyrMetThrIleThrValLeuAspThrIleSerGlnPheSer 322
Db      1532 TATTAACCAATTCGGTGAAGAAATGACACTGGCGTTTATGATGTTGATTAATCCCA 1591
Qy      323 LeuTyrAspIleLeuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleu 342
Db      1592 AATTATGACACACGACGTC-----CCAATGSAACGAAAGCA----- 1630
Qy      343 AsnGluLeuThrArgGluIleTyrThrThrGluIleAsnPheAspArgLeuProGlnIleu 362
Db      1631 ---CAACTAACAGAGAAATATAC-----GATCCTCGGGCGGGCTA 1672
Qy      363 ArgValGln-----ProAlaLeuAlaThrMetGluTyr 373
Db      1673 AACGTCTCTCAATGTTCTCGTATGACAAAGCACTCTTCCGAGATGATGAATCA 1732
Qy      374 AenLeuThrArgAlaSerPheIleuPheSerPheLeuGlnIleuIleuIleuIleuIleu 393
Db      1733 TCCGTATTTCGACCAACC---CATGATTTGATTATTAACGGGACTCAACAGTATAC 1789
Qy      394 GluAsnThrAsnIleGlyAsn---ArgLeuVal-----GlyIleSer 406
Db      1790 CATCAAGAGCACTTCTCCGCTCGCTATTAAGACATTTGGCTGCTCAAAATTAAC 1849
Qy      407 AsnArgAspAlaProThrTyrSerAsnThrIleThrGluThrLeuTyrGlyIleuIleu 426
Db      1850 TACATCGTGTGAGTGGGTACTAAT---CTTCAACAATATGATGAACTAATCA 1903
Qy      427 GlySerProThrThrIleArgProPheGluSerTyrIleValSerIleValThr 446
Db      1904 AATCTAACACGACTGTACTCTT---GATTTACAAATTAATTAACAGACTCTA 1960
Qy      447 AspArgGlnSerProIleSerProIleGlnPro-----HisPheIle 461
Db      1961 TCMAAGATGCACTACTCTTGTATTTGTTTACCTCGTTATACGATATATTTTTCGA 2020
Qy      462 IleAsnGlnIleGluLeuTyrLeuAsnGlySerSerAsnAsn-----ThrLeuIys 478
Db      2021 ATGCCAGAGTCAAGTTTTCATGTAACCAATTAATTAACGAAAGAGCTTAAG 2080
Qy      479 TyrSerAlaGlyIleSerLeuSerAlaTyrGlnAsnThrThrPheGlnPheProIle 498
Db      2081 TAT-----AATCCAGTTTCCAAAGATTAATTAACGACTACAAAG 2119
Qy      499 LysLysAspCysAsnLeuValIleAspProGlyCysSerProAsnIleAsnIleIle 518
Db      2120 GATTCGAATTAAGATTAATCTCCAGAAACTTCGATCAACCAAAATTAATGATATAC 2179
Qy      519 HisIleLeuSerHisPhe---SerLeuPheThrTyrSerTyrValIleGlyLeuGlnIle 537
Db      2180 CATAGATTATGTCATATCAACAGTATTCGCGACGGGTAAACACTACCGGATTA----- 2233
Qy      538 GlnIleLeuAspThrGlyValIleuGlyTyrThrHisSerSerValAspArgTyrAsnAla 557
Db      2234 -----GTACCTGATTTTCTTGACACATCGAAGTCAAGATTTTAAACAATTAAC 2281
Qy      558 IleSerAspIleIleThrMetIleProAlaIleIleIle---GlyAsnAsnLeuAspThr 576

```

```

Db      2282 ATATATTCAGATTAATCACTCAAAATCCCGCGTTAAATGTGGATTAATTAACGGTTT 2341
Qy      577 AsnSerIleValIleGlnGlyProGlyIleThrGlyIleAsnLeuVal---TyrLeuGln 595
Db      2342 ---GTTCCAGGTGTAAAGACACAGACATTAACGAGGGGATTTATTAAGATTAATTAAC 2398
Qy      596 SerGlnGly-----ArgLeuGlnIleThrCysGluThrPro 607
Db      2399 AGTACTGGTTCGTAGAAACCTTATTTACGTCGATATAGCGCTACCATTAAGA----- 2452
Qy      608 AenSerThrGlnSerTyrPheIleArgLeuArgTyrAlaIleThrAsnGlyIleGlyAsnThr 627
Db      2453 ---AAAGCAGGAAATATCGTGTAAAGACTGATATATGCTACTGATGCA----- 2497
Qy      628 LeuProAlaIleSerLeuThrIleProGlyValIleGlyIleProProGlnArgLeuAsn 647
Db      2498 -----GATATTGTATGATGATTAACGATGCTCAGATTCAGATTCAGATTCAGATTCAG 2551
Qy      648 AenThrPheSerGlyThrAsnTyrAsnAsnLeuGlnTyrGlyAspPheGlyTyrPheGln 667
Db      2552 CCAGGT-----GAGATCTGACATCTTAAACTTTTAAAGTTCAGAT 2593
Qy      668 PheProSerThrValThrLeuProLeuAsnArgAsnIleProPheIlePheAsnArgAla 687
Db      2594 GGTATCAACAATTAATTAATTAACAGATAGTTCGCTGATGATTAATTAATTAAGGT 2653
Qy      688 AspValSerAsnSer-----IleLeuIleIleAspIleIleGluPheIlePro 703
Db      2654 GAAGACCTTAATTAACATTAATCTGATAGTTCGTTACGTTGACCGAATCAAAATTCATCCCA 2713
Qy      704 IleThrSerSerMetHisGlnAsnArgIleuIleuIleuIleuIleuIleuIleuIleu 723
Db      2714 GTAGATGACATAT-----GAAGCGAACAAGATTTAGAACAGCAAGAAAGCA 2764
Qy      724 IleAsnThrPhePheThrAsn 730
Db      2765 GTGAATGCTGTGTTACGAT 2785

RESULT 11
US-11-108-389-17
; Sequence 17, Application US/11108389
; Publication No. US20050261188A1
GENERAL INFORMATION:
; APPLICANT: Andre R. Abad
; APPLICANT: Ronald D. Flanagan
; APPLICANT: Rafael Herrmann
; APPLICANT: Theodore W. Kahn
; APPLICANT: Albert L. Lu
; APPLICANT: Billy Fred McCutchen
; APPLICANT: James K. Presnail
; APPLICANT: James F. H. Wong
; APPLICANT: Cao-Guo Yu
TITLE OF INVENTION: Genes Encoding Proteins With Pesticidal
FILE REFERENCE: 35718/291049
CURRENT APPLICATION NUMBER: US/11/108,389
PRIORITY FILING DATE: 2003-04-18
PRIOR APPLICATION NUMBER: 60/391,786
PRIOR FILING DATE: 2002-06-26
PRIOR APPLICATION NUMBER: 60/460,787
PRIOR FILING DATE: 2003-04-04
PRIOR APPLICATION NUMBER: 10/606,320
PRIOR FILING DATE: 2003-06-25
NUMBER OF SEQ ID NOS: 134
SOFTWARE: FaestSeq for Windows Version 4.0
SEQ ID NO 17
LENGTH: 4874
TYPR: DNA
ORGANISM: Bacillus thuringiensis
FEATURE:
NAME/KEY: misc. feature
LOCATION: (0)...(0)
OTHER INFORMATION: Genomic DNA 1218-1

```

US-11-108-389-17

### Alignment Scores:

Pred. No.:	9.46e-65
Score:	675.00

Length:	4874
Matches:	316

Percent Similarity: 44.5%

Liberal:	410
Conservative:	125

**Best Local Similarity:**

**Mismatches:** 300

```

query match: 1/.46
DB: 8

```

Indels:	126
Gaps:	32

100

1

US-10-783-417-2 (1-735) x US-11-108-389-17 (1-4874)

US-10-783-417-2 (1-735) X US-11-108-389-17 (1-4874)

5 AsnAspAsnAsnGluTyrGluIleIleAspSerHisThrSerProTyrPheProAsnArg 24  
|||||  
740 AATAATCAAAATGAATATGAATTTAGATGAGCAACCTCTT-----ACTTCT 787  
25 AsnSerAsnAspSer---ArgTyrProTyrThrAsnAsnProAsnGlnProLeuGlnAsn 43  
|||||  
788 GTATCAATGATTTCTAACAGATACCCCTTTGCGAATGAGCCAAACAAATGCGCTACAAAT 847  
44 ThrAsnTyrLysGluTyrLeuAsnMetCysGlnGlyAsn---ThrGlnTyrGlyAspAsn 62  
|||||  
848 ATGATATTAAAGATTATTAAATAATGCTCGCGGAATGCTGAGGAATACCTCGTTCA 907  
63 PheGluThrPheAlaSerAlaAspThrIleAlaAlaValSerAlaGlyThrIleValSer 82  
|||||  
908 CCTGAGTAAGTCTTGTAACGAGCAAAAGAT-----CGAGCTAAGGCGCGCAATTGATGATGA 961  
GlyThrLeuLeuAlaGlyIleGlyIleGlyLeuThrSerIleSerGlyProIleGlyIleIle 102  
|||||  
962 GGTAAATTAATCACTACAGGTTAGGG-----GTCCCATTTATT 997  
103 GlyAlaIleIleIleSerPheGlyThrLeuIleThrValPheTyrProAlaGlyGluGln 122  
|||||  
998 GGGCCGATAGTACAGCTTATTACTCACTATTATGATATTCTGTGCGCTTCACAGGAA-- 105  
123 AspTyrThrValTyrThrGlnPheIleIleMetGlyIleIlePheValAspThrProLeu 142  
1055 ---AAGAGTCATGCGGAATTTTATTTGCAACAGTGAAGAACTCATTAATCAAAAATA 1111  
143 ThrGluSerIleGlyGlnLeuLysLeuGlnThrIleGlnGlyIlePheArgGlnIleLeuGln 162  
|||||  
1112 GCAGATATGCAAGGAATTAAGCGCTTCGCGAATTGAAGGATTGAGTAAATTAATCAAA 1177  
163 SerTyrAsnThrIleAlaLeuAspAspTyrArgLysLeuLysArgLeuGlnAlaIleProLysLeu 182  
1172 TTATATCTAATCAGCGCTTGAAGAAATGGGAGA-- 120  
183 ProProSerSerAlaLeuGlnGlnAlaAlaLeuThrLeuLysIleArgPheGluAsnVal 202  
1205 AATCCAAATGGTTCA-----AGACCCCTTACAGAGATGTCGCAAAATCGAATTTGAAATCCG 1258  
203 HisAsnAspPheIleArgGluIleProGlyPheGlnLeuGlnIleThrTyrIleThrIleLeuLeu 222  
1259 GATAGCTTATTATTCGCAATATATGCGATCTTTTAAAGTACAAATTTGAAAGTACCAATTC 1318  
223 LeuProIleTyrAlaGlnAlaAlaAsnPheHisLeuAsnLeuLeuGlnGlnIleValGlu 242  
1319 CTTACTGTAATGCAATGACGACGCCAACCTTACTTACGTATTATTAAGACCGCGTCAAT 1378  
243 LeuAlaAspGluTyrAsnAlaAspIleHisProSerGlnIleGluProAsnAlaGlyThr 262  
1379 TTGGGAGGAATGGGGATGG-----TCAACAATCTACT 1411  
263 SerAspAspTyrTyrLysLeuLeuLysGluAsnIleProLysTyrSerAsnTyrCysAla 282  
1412 ATTAATTAATCTATTATGATGTCGCAATGAAACTTACTGCAAGAAATATTCGATACAGTGTA 1477  
283 AsnTyrTyrArgThrGlyLeuLysAsnLeuArgAspGluProAsnMetCysTyrSerIle 302  
1472 AAGTGTATGAACCTGTTAGCAAAATTTAAAGCAGACGCGCTTAACAAATGGGTTGAC 1533  
303 PheAsnAspTyrArgArgTyrMetThrIleThrValLeuAspThrIleSerGlnPheSer 322

Db 1532 TATACCAATTCCTCGTAGGAATGACACTGGCGGTATTAAATGTTGTCGATTAATTCCTCA 1591

Qy 323 LeuTYRApIleuTYRARGTYRARGApSerTleGlyGlyIleGluVallySgIlyIleYs 342

Db 1592 AATTATGACACACGACGCTAC-----CCAATGGAAACGAAGA----- 1630

Qy 343 AenGluLeuThrArgGluIleTYRThrThGluIleasnPheAspArgLeuProGlnIleu 362

Db 1631 ---CAACTTAACAGGAAGATATATACA-----GATCCACTGGCGCGGTGA 1672

Qy 363 ArgValGln-----ProAsnLeuAlaThrMetGluTYR 373

Db 1673 AACGTGCTTCAATYGTTGGTCTCGTAGTACAAAGACACTTCTTCCGAGATGATGAAATCA 1732

Qy 374 AsnLeuThrArgAlaSerPheYsLeuPheSerPheLeuGlnPheIlePheTYRThr 393

Db 1733 TCCGTTATGACACACACC-----CAGTATTTGATTATATACGGGACTCACAGGTATATACA 1789

Qy 394 GluAsnThrAsnPheGlyAsn--ArgLeuVal-----GlyIleSer 406

Db 1790 CAATCAAGAAGCATTTCTCGCTCGCTGTATATATAGACATTGGCGGTGTCATCAATTAAGC 1849

Qy 407 AsnArgAspAlaProThrTYRYSerAsnThrIleThrGluThrLeuTYRglyValArgThr 426

Db 1850 TACCATGCTGCTAGTAGGGGTAGTAAT-----CTTCAACAAATGATGGAACATAATCA 1903

Qy 427 GlySerProThrTYRThrTYRThrIleArgProPheGluSerTYRlySvalSerIleValThr 446

Db 1904 AATCTACACAGCATCTAGTACCTTT---GATTTTACGAATTTATGATTTATCAAGACTCTA 1960

Qy 447 AspArgGlnSerProProValSerProIleGlnPro-----HisPheIle 461

Db 1961 TCAAAGAGTACAGTACTCTGATATATGTTTACCTCGTTATACGTATATATTTTGTGA 2020

Qy 462 IleAsnGlnIleGluLeuTYRleuAsnGlySerSerAsnAsn-----ThrIleYs 478

Db 2021 ATGCCAGAAAGTCAGGTTTTCATGTGTAAACCAATTGAATATATACAGAAAGCGTTTAAG 2080

Qy 479 TYRserAlaGlyGlySerLeuSerAsnTYRglnAsnThrThrPhePheGlnPheProArg 498

Db 2081 TAT-----AATCCAGTTTCCAAAGATATATATAGGAGATACAGA 2119

Qy 499 LysIleAspCyAsnLeuValIleAspProGlyCysSerProAsnPheAsnTYRser 518

Db 2120 GATTCGGAATTAGATTAATCTCCGAAACTTCAGATCAACCAATTAATGTAGCTATATAGC 2179

Qy 519 HisIleLeuSerHisPhe---SerLeuPheThrTYRserTYRValIleGlyIleuGlnLeu 537

Db 2180 CATGAGATTATGTCAATATCAACAGTAATCCCGCGAGCGGTAACTACCGGATTA----- 2233

Qy 538 GlnIleLeuAspThrGlyValIleuGlyTYRThrHisSerSerValIleAspTYRAsnAla 557

Db 2234 -----GACTCGTATTTTCTTGACACATCGAAGTGCAGATTTTAAACATATCA 2281

Qy 558 IleSerAspYsIleIleThrMetIleAspAlaIleYs--GlyAsnAsnLeuAspThr 576

Db 2282 ATATATTCAGATTAATAATCACTCAATATCCGCGCGTTAAATGTGGAGTAATTTTACCGTT 2341

Qy 577 AsnSerIysValIleGlyGlyProGlyHisThrGlyIysAsnLeuVal---TYRleuGln 595

Db 2342 ---GTTCCAGGTGTAAAGACCAAGCAATACAGAGAGGAGATTAATTAACGTATATATAGA 2398

Qy 596 SerGlnGly-----ArgLeuGluIleThrCysGluThrPro 607

Db 2399 AGTACTGTTCTGTAGGAACTTATTTCTAGCTCGATATAGCGCTTACGACTTATGAA----- 2452

Qy 608 AsnSerThrGlnSerTYRpheIleArgLeuArgTYRAlaThrAsnGlyAlaGlyAsnThr 627

Db 2453 ---AAAGAGGGAATATTCGTGTAAAGCTAGATATCTACTATGCA----- 2497

Qy 628 LeuProAsnIleSerLeuThrIlePheProGlyValIleGlyIleProGlnAsnGlyLeuAsn 647

```

Db      2498 -----GATATTGATTCAGTGAACGATGCTCAGATTCAAGTCCAAAACATGAAC 2551
QY      648 AanthrPhseSerGlyThrAntyRhaAnaenleuGlnTyGlyAspPheGlyTyRheGln 667
Db      2552 CCAAGT-----GAGGATCTACATCTTAAACCTTTAAAGTTGCAGAT 2593
QY      668 PheProSerThrValThrleuProleuAsnArgAsnIleProPheIlePheAnaGala 687
Db      2594 GGTATACACAACATTAAATTATAGACAGATAGTCCCTAGCATGAAACATAATTAGGT 2653
QY      688 AspValSerAsnSer-----IleLeuIleIleAspLysIleGluPheIlePro 703
Db      2654 GAAGACCCATTAATTCATTAATCTGATATAGTTAGCTGACCGAATCGAATTCATCCCA 2713
QY      704 IleThrSerSerMetHisGlnAnaenArgGluLysGlnLysleuGlnThrIleGlnThrLys 723
Db      2714 GTTAGATGACACATAT-----GAAGCGAACAAGATTATAGACGACGAGAAAGCA 2764
QY      724 IleAanthrPhePheThrAsn 730
Db      2765 GTGAATGCCTTGTTTACGAAAT 2785

RESULT 12
US-11-091-643-3
; Sequence 3, Application US/11091643
; Publication No. US20050246789A1
; GENERAL INFORMATION:
; APPLICANT: TANAKA, Masao
; APPLICANT: YOKOYAMA, Tomoko
; APPLICANT: AOYAGI, Moriochi
; APPLICANT: HASEGAWA, Makoto
; APPLICANT: EHARA, Gaku
; APPLICANT: KIMURA, Masaharu
; APPLICANT: NISHIHASHI, Hideji
; TITLE OF INVENTION: Polypeptide having larvae growth inhibiting or
; TITLE OF INVENTION: insecticidal effect on scarabaeidae insects and
; FILE REFERENCE: OPI335
; CURRENT APPLICATION NUMBER: US/11/091,643
; CURRENT FILING DATE: 2005-03-29
; PRIOR APPLICATION NUMBER: JP 2001-115754
; PRIOR FILING DATE: 2001-04-13
; PRIOR APPLICATION NUMBER: JP 2001-203463
; PRIOR FILING DATE: 2001-07-04
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 4359
; TYPE: DNA
; ORGANISM: Bacillus popilliae
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (282)..(4229)
US-11-091-643-3

Alignment Scores:
Pred. No.:      3,78e-64      Length:      4359
Percent:        669.00      Matches:      230
Score Similarity: 43.5%      Conservative: 106
Best Local Similarity: 29.8%      Mismatches:   301
Query Match:    17.3%      Indels:      136
DB:            10      Gaps:      34

US-10-783-417-2 (1-735) x US-11-091-643-3 (1-4359)
QY      1 MetAsnGln-----AsnAsnAspAsn-----AengLufyr 10
Db      336 ATCAATCAATATATATACCAAAACGATTAACAAAAGTTACACCAAAAGTGAAATGAATG 395
QY      11 GluIleIle-----AspSerHisThrSerProTyR-PheProAsnArgAsnSerAsnAsp 28
Db      396 CAATCATTCACACCTTCAAGTAAAGCTTTACTTACAGTCCCAAT-----440

```

```

QY      29 SerArgTyR-ProTyR-ThrAsnAsnProAengGlnProLeuGlnAsnThrAsnTyR-LysGlu 48
Db      441 ---AGTATCCGATATCTACGATCCGATTCATGATGATGACAGAGGCGTGAAGTTATTAATAAT 497
QY      49 TrpLeuAsnMetCysGlnGlnLysAnthrGlnTyRgLyAspAsnPheGluThrPheAlaSer 68
Db      498 TGGTGGATATATGTATACAGGACA-----GGCGACACACGAAGCCCGGAACCTGCT 548
QY      69 AlaAspThrIleAlaAlaValSerAlaGlyThrIleValSerGlyThrLeuLeuAlaGly 88
Db      549 GCTATTTCAAAAGGCTGTCTTCTGCTGCAATTACTATTAAC-----590
QY      89 IleGlyLysLeuThrSerIleSerGlyProIleGlyIle-----IleGly 103
Db      591 -----ACCGGCTTCTGCTTACTAGAGTGTCCGTTCATCACAACAATCCGG 638
QY      104 AlaIleIleIleSerPheGlyThrleuIleThrValPheTyRProAlaGlyLysGluAsp 123
Db      639 GCATTT-----TATACCTTCTATTAAGAAATACCTTAATGCTTCA-----AGC 680
QY      124 LysThrValTrpThrGlnPheIleLysMetGlyLysIlePheValAspThrProLeuThr 143
Db      681 AATACTCAATGGAGAGATTATAGCACATGTGGAAACCTCATTAATGCAGAAACCTACACA 740
QY      144 GluSerIleLysGlnLeuLysleuGlnThrleuGlnGlyPheArgGlnIleLeuGlnSer 163
Db      741 GATCATGTAGAATTCGGCACTTACAAATTAATGCTTTACGTAAATACATAGAGATA 800
QY      164 TyRAsnThrAlaLeuAspAspTrpArgLysLeuLysArgLeuGlnAlaProGlyLeuPro 183
Db      801 TATAACGAAGCTTTATATAGTTGGAACAA-----GATCCT 836
QY      184 ProSerSerAlaLeuGlnGlnAlaAlaLeuThrleuLysIleArgPheGluAsnValHis 203
Db      837 AACAAATGCAAACTTAAGAT-----GATGTAGAAGTAATTCGTAGGCTCAAT 887
QY      204 AsnAspPheIleArgGluIleProGlyPheGlnLeuGlnThrTyRAsnThrLeuLeuLeu 223
Db      888 AGCCAATTCGAAGATATATTCACAAATTTAAAGAAAGGTTTGAAGTTCAATTATTA 947
QY      224 ProIleTyRAlaGlnAlaAlaAsnPheHisLeuAsnLeuGlnGlnGlnAlaGluLeu 243
Db      948 ACTATATATGACAAATCTGCAATCTTCATCTATATATTAAGAGATTCCTCTTGAT 1007
QY      244 AlaAspGluTrpAsnAlaAspIleHisProSerGlnIleGluProAsnAlaGlyThrSer 263
Db      1008 GGTGCATCTTGGGATTT-----GCTCAAGCTACTATT 1040
QY      264 AspAspTyR-TyrLysLeuLeuLysGluAsnIleProLysTyRSerAntyRysAlaAsn 283
Db      1041 GACATATATTAATCAATGCCCAATTAAGAAACCGCAGAGTATGCATATCTATACCACT 1100
QY      284 ThrTyRArgThrGlyLeuLysAsnLeuArgAspGluProAsnMetLysTrpSerIlePhe 303
Db      1101 TGGTATCAGACCGGTTTACAAAGATTCGAAGGACACTACGTACAGCTTGCTCTGTAT 1160
QY      304 AsnAspTyRArgArgTyRMetThrIleThrValLeuAspThrIleSerGlnPheSerLeu 323
Db      1161 CATGATTTTAAAGAAAGAAATGACACTTACAGATTAATGATTCCTCATTAATTTCAAT 1220
QY      324 TyRAspIleLysArgTyRArgAspSerIleGlyLysIleGlyValLysGlyIleLysAsn 343
Db      1221 TATGATGCCGTAAGTTAC-----CCACTGAGGTAAAGGA-----1256
QY      344 GluLeuThrArgGluIleTyRThrThrGluIleAsn-----PheAspArg 358
Db      1257 GAGCTTAGAGAGAAATTTATACGATTCAGATGACCCGATTAATCGAATGATGCA 1316
QY      359 LeuProGlnLeuArgValGlnProAsnLeuAla---ThrMetGlyTyRAsnLeuThrArg 377
Db      1317 GCACCATCATTCGCAGAAATTAAGAAATCTTAAGATTAGGCAACGAAGCTGTACTCTGG 1376
QY      378 AlaSerPheLysLeuPheSerPheLeuGlnPheIlePheTyRThrGluAsnThrAsn 397

```

Db 1377 ATATCCGGTGAATTATAGTATATACAGGTAGATTGTACGGCTATACGTATATACGAT 1436  
Phe-----GlyAsnArgLeuValGlyIleSerAsnArg-----AspAlaPro 411  
1437 TATTGGGCGACACACAGGCTAGATTCTTGAACCAATGTTATCGGTTTACAGGCTCT 1496  
412 ThrTyrSerAsnThrIleThrGluThrLeuTyrGlyGluArgThrGlySerProThrThr 431  
1497 ACCTATGATGACACGATTAATATAGT-----CGTACGATTTCTATCCGANG 1544  
432 LysThrIleArgProPheGluSerTyrLysValSerIleValIleThrAspArgIleSerPro 451  
1545 AATTCTATT-----GATGTTTATTCCTACTACGTAGTACCTGTTGCTCTCGTGG 1595  
452 ProValSerProIleGluProHisPheIleIleAsnGluIleGluLeuTyrLeuAsnGly 471  
1596 CCAACTGGC-----GGTTTGTGTTGGAGAGTGGCTGGCTGATTTTTCG 1643  
472 SerSerAsnAsnThrLeuLysTyrSerAlaGlySerLeuSerAsn-----TyrGln 489  
1644 AAAAGTCTAGCAC-----GGTTTATTTAGTGAGCGGGGTATCAG 1685  
490 AsnThrThrPhePheGluPheProArgLysLysAspCysAsnLeuValIleAsp--Pro 508  
1686 AATCCAGTATATATTT-----TCGAGTTCCACTTATTACTTTTAACTTAACTT 1730  
509 GlyCys-----SerProAsnPheAsnAsnTyrSerHisIleLeuSerHisPheSer 525  
1731 GGAGTAGACCAAGATACGCCAATGCTGCCGACTATAGTATACATTAACGTATGATCACA 1790  
526 LeuPheThrTyrSerTyrValIleGlyLeuGluIleLeuAspThrGlyValLeu 545  
1791 GCATTTTCCA-----ACTGGATTGAATGGAAGCTGT-----CCGGTTT 1829  
546 GlyTyrPheHisSerSerValAspArgTyrAsnAlaIleSerAspLysIleIleThrMet 565  
1830 GGACGGTATTCGCAACTGTTAGTCCGTCGACAAATGAAATGAGCCAGCAAAATTAACGCAA 1889  
566 IleProAlaIleLeuGlyAsnAsnLeuAspThrAsnSerLysValIleGluGlyProGly 585  
1890 ATCCCGGCTGTATGACAACTCCCTCGAC--AATTGCCAGTATTTAGAGGAGCTGGA 1946  
586 HisThrGlyGlyAsnLeuValTyrLeuGluIleSerGluIleArgLeuGluIleThrCysGlu 605  
1947 TTACACGAGGAGAGCTGTTGAAGACAGTATCTTAGTGTGTTTGTCTTAATCACT 2006  
606 ThrProAsnSerThrGlnSerTyrPheIleArgLeuArgTyrAlaThr----- 621  
2007 TCATCGAGGCGGCGCAATCTTACCGCATCCGGTCCGTTATGCTGCTCGAGTAGATTGA 2066  
622 ---AsnGlyAlaGlyAsnThrLeuProAsnIleSerLeuThrIleProGlyValIleGly 640  
2067 ATTATGAGTATATTTCTATATGACCTTTTATTTCCAAAGAAATTAAGTCTTACCAATCA 2126  
641 IleProProGluArgLeuAsnAsnThrPheSerGlyThrAsnTyrAsnAsnLeuGluIle 660  
2127 ATGCCACCA-----CTGACCGAAACT-----GTACCTTACGAACCTTTTAAATTT 2171  
661 GlyAspPheGlyTyrPheGluPheProSerThrValIleLeuProLeuAsnArgAsnIle 680  
2172 GCAGATTTTGGTGTGTC-----ACCTTTTAAACAAGCTACTGCTTAAACAAAGATAT 2219  
681 ProPheIlePheAsnArgAlaAspValSerAsnSerIleLeuIleIleAspLysIleGlu 700  
2220 ACTTTTCAATTCAT-----ACGGGTGACGACGCTATATTAATTAACGAAATTTGA 2267  
701 PheIleProIleThrSerSerMetHisGluAsnArgGluLysGluIleLeuGluIleThrIle 720  
2268 TTTGTTCCAAATTTGAGGAGTGTGTTTTCAGAGTACGAAACCAACACAGCTTAAAGAAAAGCA 2327  
721 GlnThrLysIleAsnThrPhePheThrAsnHisThrLys 733

Db 2328 AGGAAAGCGGTGAACCAATTGTTTACAGATGATCGAAA 2366  
RESULT 13  
US-11-192-801-3  
/ Sequence 3, Application US/11192801  
/ Publication No. US20050273882A1  
/ GENERAL INFORMATION:  
/ APPLICANT: Romano, Charles P.  
/ TITLE OF INVENTION: Improved Expression of Cry3Bb Insecticidal Protein in Plants  
/ FILE REFERENCE: 38-21(15304) Cry3Bb Improved Exp. Corn  
/ CURRENT APPLICATION NUMBER: US/11/192,801  
/ CURRENT FILING DATE: 2005-07-29  
/ PRIOR APPLICATION NUMBER: US/10/232,665  
/ PRIOR FILING DATE: 2002-08-29  
/ PRIOR APPLICATION NUMBER: US/09/377,466  
/ PRIOR FILING DATE: 1999-08-19  
/ NUMBER OF SEQ ID NOS: 43  
/ SOFTWARE: PatentIn Ver. 2.0  
/ SEQ ID NO 3  
/ LENGTH: 1959  
/ TYPE: DNA  
/ ORGANISM: Bacillus thuringiensis  
/ FEATURE:  
/ NAME/KEY: CDS  
/ LOCATION: (1)..(1956)  
/ OTHER INFORMATION: naturally occurring nucleotide sequence encoding a  
/ OTHER INFORMATION: Cry3Bb2 amino acid sequence  
US-11-192-801-3  
Alignment Scores:  
Pred. No.: 1,066-63 Length: 1959  
Score: 660.00 Matches: 209  
Percent Similarity: 43.0% Conservative: 111  
Best Local Similarity: 28.1% Mismatches: 290  
Query Match: 17.1% Indels: 134  
Gaps: 29  
US-10-783-417-2 (1-735) x US-11-192-801-3 (1-1959)  
QY 1 MetLangInAsnAsnAspAsnAsnGluTyrGluIleIleAspSerHisThrSerProTyr 20  
Db 1 ATGAATCCAAACAT--CGAAGTGAACATGATACGATTAAGTTACACCTTAACAGTGA 57  
QY 21 PheProAsnArgAsnSerAsnAspSerArgTyrProTyrThrAsnAsnProAsnGluPro 40  
Db 58 TTGCCA-----ACTAACCATATCATATCTTAGCTGACGACATCCAAATTCGACA 108  
QY 41 LeuGluAsnThrAsnTyrLysGluTyrLeuAsnMetCysGlu--GlyAsnThrGluIle 59  
Db 109 CTAGAAAGATTAATAATTATTAAGAAATTTTAAAGAAAGACGAGACAGCTTACGGAAGTG 168  
QY 60 GlyAspAsnPheGluThrPheAlaSerAlaAspThrIleAla--AlaValSerAlaGly 78  
Db 169 CTAGCAAC-----CTTACGTAAGAAAGTACGACTTGGACAGCA 207  
QY 79 ThrIleValSerGlyThrLeuLeuAlaGlyIleGlyGlyLeuThrSerIleSerGlyPro 98  
Db 208 ATTCGTGTTGAGGCAAGATTTTAGGTGTGTAAGA----- 243  
QY 99 IleGlyIleIleGlyAlaIleIleIleSerPheGlyThrLeuIleThrValPheTyrPro 118  
Db 244 GTTCATTTGCTGGGCGACTCACTCATTTTATTCATCATCTTTCTTGACACTATATGCGCA 303  
QY 119 AlaGlyGluGluAspLysThrValIleThrGluPheIleLysMetGlyGluIlePheVal 138  
Db 304 AGT-----GATGCTGACCCATGAAAGCTTTTATGCGCAAGTTGAAGTACTGATA 354  
QY 139 AspThrProLeuThrGluSerIleLysGluLeuLysValLeuGluThrLeuGluGlyPheArg 158  
Db 355 GATTAAGAAATTAAGAGATATCTTAAGGTAAAGCTCTTGACAGAGTTCAGAGGTCTTCA 414  
QY 159 GlnIleLeuGluSerTyrThrAsnThrAlaLeuAspAspTPAArgLysLeuLysArgLeuGlu 178

Db 415 AATAATTTGGAAGATTATGTTAAATGCGTTAAATTCCTGGAGAA----- 459  
 Qy 179 AlaProGlyLeuProSerSerAlaLeuGlnAlaLeuThrLeuYsIleArg 198  
 Db 460 -----ACACCTTTAAGTTTGGCAAGTAAAGAACAGATGGAATGGAACCT 510  
 Qy 199 PheGluAsnValHisAsnAspPheIleArgGluIleProGlyPheGlnLeuGluThrTy 218  
 Db 511 TTTTCACAGCAGAAAGTCAATTTTGGTAATTCATGCCGTCATTTGCAAGTTTCCAAATTC 570  
 Qy 219 LysThrLeuLeuLeuProIleTyAlaGlnAlaAlaAsnPheHisLeuAsnLeuGln 238  
 Db 571 GAAGTCGTGTTTCTACCAATATGCAACAGCTGCAAAATACATTAATGCTTAATTAATAA 630  
 Qy 239 GlnGlyAlaGlnLeuAlaAspGluThrPheAlaAspIleHisProSerGlnIleGluPro 258  
 Db 631 GATGCTCAAGTTTGTGAGAAAGATGGGATATCT----- 666  
 Qy 259 AsnAlaGlyThrSerAspAspTyTyTyLeuLeuLeuGlnLeuAsnIleProLysTySer 278  
 Db 667 TCAGGAAGATGTTCTGCAATTTTATCATAGCAATTAACCTTACG---CAACATATCACT 723  
 Qy 279 AsnTyTyCysAlaAsnThrTyArgThrGlyLeuLysAsnLeuArgAspGluProAsnMet 298  
 Db 724 GACCATTTGTGCAATTTGATATATGTTGATTAATGTTTAAGAGTTCAACTTATGAT 783  
 Qy 299 LysIlePheSerIlePheAsnAspTyTyArgGlyTyMetThrIleThrValLeuAspThrIle 318  
 Db 784 GCATGGGTCAAAATTTAACGTTTTCGACAGAAATACCTTAATCTATTAATGATCTTAAT 843  
 Qy 319 SerGlnPheSerLeuTyTyAspIleLysArgTyTyArgAspSerIleGlyGlyIleGluVal 338  
 Db 844 GATCTTTCCCATTTATGATGTTGCGTTACTCA----- 879  
 Qy 339 LysGlyIleLysAsnGlnLeuThrArgGluIleTyThrThrGlnIleAsnPheAspArg 358  
 Db 880 AAAGGTGTAAACACAGACTTAACAGAGACATTTTACGGATCAATY---TTTTCACATC 936  
 Qy 359 LeuProGlnLeuArgValGlnProAsnLeuAlaThrMetGluTyTyAsnLeuThrArgAla 378  
 Db 937 AATACTCTTCAGAGATGAGCAACACTTTTGTGATAGTAAACCTTATTCAGAAACCT 996  
 Qy 379 SerPheLysLeuPheSerPheLeuGlnGlnPheIlePheTyThrGlnLeuThrAsnPhe 398  
 Db 997 -----CATTAATTTGATTAATTTACAGCGTATTGAATTCATACG----- 1035  
 Qy 399 GlyAsnArgLeu---ValGlyIleSerAsnArgAspAlaProThrTy----- 413  
 Db 1036 -----CGTCTTCAACCTGTTACTCTGGAGAAAGATTTCTTCAATTAATGCTGTAAT 1089  
 Qy 414 -----SerAsnThrIleThrGlnThrLeuTyGly 423  
 Db 1090 TATGTAGAAACTAGACCTAGTATAGATCATGTAACATCAATCTCCCATTTATGGA 1149  
 Qy 424 GluArgThrGlySerProThrThrLysThrIleArgProPheGlnSer-----Ty 440  
 Db 1150 GATTAATCTACTGACACCTGTACAAAGTTA-----AGCTTGTAGACAAAGAAAGTTTAT 1203  
 Qy 441 LysValSerIleValThrAspArgGlnSerProProValSerProIleGlnProHisPhe 460  
 Db 1204 CGAACCTATAGCTAATACAGACGAGCGGCTTGCCGCAATGCGCAAGTA-----TATTTT 1257  
 Qy 461 IleIleAsnGlnIleGluLeu-----TyLeuAsnGlnLysSerSerAsnAsn 475  
 Db 1258 GGTGTATACGAAGTATTTTATGTAATATGATGATCAAAAGAAATGAACCTAGTACAA 1317  
 Qy 476 ThrLeuLysTySerIleGlyLysSerLeuSerAsnTyGlnAsnThrThrPhePheGln 495  
 Db 1318 ACATATGATTCAGAAAGAAAGAAATGGCATGTAGTGCGACAGAGATTTCTATGCAATTA 1377  
 Qy 496 PheProArgLysLysAspCysAsnLeuValIleAspProGlyCysSerProAsnPheAsn 515  
 Db 1378 CACCGAGAAACACACAGT-----GACACCTTGAAGAA 1410

Qy 516 AsnTyTySerHisIleLeuSerHisPheSerLeuPheThrTySerTyValIleGlyLeu 535  
 Db 1411 GCATATAGTCATCAGCTTAATTAACGGGAAATGTTCTTA-----ATG 1452  
 Qy 536 GlnLeuGlnIleLeuAspThrGlyValLeuGlyTyTyPheHisSerSerValAspArgTy 555  
 Db 1453 CAGAACCGTCGGAACCAATTCATTTTATCTTGGACATATGAAGTGTAGCTTTT 1512  
 Qy 556 AsnAlaIleSerAspLysIleIleThrMetIleProAlaIleLysGlyAsnLeuAsp 575  
 Db 1513 AATACATTTGATGCTGAAAGATTAATCTCACTTCAGTAGTAACCATATGCTTGCT 1572  
 Qy 576 ThrAsnSerLysValIleGlnGlyProGlyHisIleThrGlyLysAsnLeuValTyLeuGln 595  
 Db 1573 TCAGGTCTTCATTTATTAAGGTCAGGATTCACAGGAGAAATTAATCTATTCCTAAA 1632  
 Qy 596 -----SerGlnLysArgLeuGlnIleThrCysGluThrProAsnSerThrGln 611  
 Db 1633 GAATCTAGTAATTCATTCCTAAATTTAAAGTTACATTAATTCAGACCTTGTTACA 1692  
 Qy 612 SerTyTyPheIleArgLeuArgTyTyAlaThrAsnGlyAlaGlyAsnThrLeuProAsnIle 631  
 Db 1693 CGATATCGTGAAGATATAGCTATGCTTCTAC-----ACTAATCTTA 1734  
 Qy 632 SerLeuThrIleProGlyValIleGlyIleProProGlnArgLeuAsnAsnThrPheSer 651  
 Db 1735 CAGCTTTTGTG-----CAAAATTCAAACAAATGATTTAT 1770  
 Qy 652 GlyThrAsnTyAsn-----AsnLeuGlnTyGlyAspPheGlyTyTyPheGlnPhe 668  
 Db 1771 GTCATCTACATTAATTAACCTATGATATAGT---GAGATTTAACATATCAACATTT 1827  
 Qy 669 ProSerThrValThrLeuProLeuAsnArgAsnIleProPhe----- 682  
 Db 1828 GATCTCGCAACTACT-----AATCTAATATGCGGTCTCGGTGATACGAATGA 1878  
 Qy 683 -----IlePheAsnArgAlaAspValSerAsnSerIleLeuIleIleAspLysIleGln 700  
 Db 1879 CTTATATATGAGCGACGAATCTTGTCTTAATGAAAAATCATATATGATATGATGAA 1938  
 Qy 701 PheIleProIle 704  
 Db 1939 TTTATCCAGTA 1950  
 RESULT 14  
 US-11-058-727-73  
 ; Sequence 73, Application US/11058727  
 ; Publication No. US20050261483A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Andre R. Abad  
 ; APPLICANT: Ronald D. Flannagan  
 ; APPLICANT: Rafael Herrmann  
 ; APPLICANT: Theodore W. Kahn  
 ; APPLICANT: Albert L. Lu  
 ; APPLICANT: Billy Fred McClutchen  
 ; APPLICANT: James K. Presnail  
 ; APPLICANT: James F. H. Wong  
 ; APPLICANT: Cao-Guo Yu  
 ; TITLE OF INVENTION: Genes Encoding Proteins With Pesticidal  
 ; TITLE OF INVENTION: Activity  
 ; FILE REFERENCE: 35718/287809  
 ; CURRENT APPLICATION NUMBER: US/11/058,727  
 ; PRIOR FILING DATE: 2005-02-15  
 ; PRIOR APPLICATION NUMBER: 60/391,786  
 ; PRIOR FILING DATE: 2002-06-26  
 ; PRIOR APPLICATION NUMBER: 60/460,787  
 ; PRIOR FILING DATE: 2003-04-04  
 ; PRIOR APPLICATION NUMBER: 10/606,320  
 ; NUMBER OF SEQ ID NOS: 134  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 73





QY 374 AenLeuThrArgAlaSerPheLeuPheSerPheLeuGluGlnPheIlePheTyrThr 393  
Db 1021 TCCGTTATTGACCAACCC---CATGATATTGATTATATACCGGACCTCAGCTGATATCA 1077  
QY 394 GluAenThrAsnPhedGlyAsn---ArgLeuVal-----GlyIleSer 406  
Db 1078 CATCAAGAGATTCCTTCGCGCTATATATGAACATTTGGCTGTCATCAATAAAGC 1137  
QY 407 AenArgAspAlaProThrTyrSerAenThrIleThrGluThrLeuTyrGlyIuArgThr 426  
Db 1138 TACCATCGTGTGATGAGGGGTATAT-----CTTCAACAATGATGAACTATCA 1191  
QY 427 GlySerProThrThrIleThrIleArgProPheGluSerTyrIleValSerIleValThr 446  
Db 1192 AATCTACACAGCACTAGTACCTT---GATTTACCAATATGATATTTACAAGACTCTA 1248  
QY 447 AspArgGlnSerProProValSerProIleGlnPro-----HisPheIle 461  
Db 1249 TCAAAGATGACGATCTCTGATATATTGTTTACCTGGTTATACGTATATATTTTGGGA 1308  
QY 462 IleAenGlnIleGluLeuThrLeuAenGlySerSerAsnAsn-----ThrLeuIys 478  
Db 1309 ATGCCAAGAGTCAAGTTTTCATGTGAACCAATTGAATATACCAAGAAAGCTTAAG 1368  
QY 479 TyrSerAlaGlyGlySerLeuSerAenTyrGlnAenThrThrPhePheGlnPheProArg 498  
Db 1369 TAT-----AATCAGTTTCCAAGATATTTATAGCAGTACAAGA 1407  
QY 499 IysIysAspCysAenLeuValIleAspProGlyCysSerProAenPheAenThrTyrSer 518  
Db 1408 GATTCGGAATTAGAACTTACCTCCAGAACTTCAGATCAACAAATTATGATCATATAGC 1467  
QY 519 HisIleLeuSerHisPhe---SerLeuPheThrTyrSerTyrValIleGlyLeuGlnLeu 537  
Db 1468 CATAGATTATGTCATATCAAGATATCCCGGACCGGTAAACACTACCGGATTA----- 1521  
QY 538 GlnIleLeuAspThrGlyValLeuGlyTyrPheHisSerSerValAspArgTyrAsnAla 557  
Db 1522 -----GTACCTGATATTTCTTGACACATCGAAGTGCAGATTTTAAACAATACA 1569  
QY 558 IleSerAspIysIleIleThrMetIleProAlaIleIys---GlyAsnAenLeuAspThr 576  
Db 1570 ATATATTCAGATAAATCACTCAAAATTCGCGCGTTAAATGTTGGATTAATTACCGTTT 1629  
QY 577 AenSerIysValIleGluGlyProGlyHisThrGlyGlyAsnLeuVal---TyrLeuGln 595  
Db 1630 ---GTTCCAGTGTAAAGACCAAGACATACAGAGGGGATTTATACGTAATAATAGA 1686  
QY 596 SerGlnGly-----ArgLeuGlnIleThrCysGluThrPro 607  
Db 1687 AGTACTGCTGTGTAGGAACCTTATTTCTAGCTCGATATGCGCTAGCATTAAGAA----- 1740  
QY 608 AenSerThrGlnSerTyrPheIleArgLeuArgTyrAlaThrAsnGlyAlaGlyAenThr 627  
Db 1741 ---AAAGCAGGAAATATCGTATAGACTGAGATATGCTACTAGTCA----- 1785  
QY 628 LeuProAenIleSerLeuThrIleProGlyValIleGlyIleProGlnArgLeuAen 647  
Db 1786 -----GATATTGTATTGATGATGAACGATGCTCAGATTCAGATGCCAATAATGAAAC 1839  
QY 648 AenThrPheSerGlyThrAenTyrAenAenLeuGlnTyrGlyAspPheGlyTyrPheGln 667  
Db 1840 CCAGGT-----GAGGATCTGACATCTAAACCTTAAAGTTGCAGAT 1881  
QY 668 PheProSerThrValThrLeuProLeuAenArgAenIleProPheIlePheAenArgAla 687  
Db 1882 GCTATCAACAACAGTTAATTAGCAACAGTAGTTCGTTGCAGTTAAACATTAATTAGGT 1941  
QY 688 AspValSerAsnSer-----IleLeuIleIleAspIysIleGluPheIlePro 703  
Db 1942 GAAGACCTTAATTCAATTATCTGTATGTTAGTTTACGTTGACCGAATCGAATTATCATCCA 2001  
QY 704 Ile 704

Db 2002 GTA 2004

Search completed: January 20, 2006, 23:03:36  
Job time : 382 secs



GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - nucleic search, using frame\_plus.p2n model

Run on: January 20, 2006, 17:27:41 ; Search time 1439 Seconds  
(without alignments)  
4223.762 Million cell updates/sec

Title: US-10-783-417-2  
Perfect score: 3869  
Sequence: 1 MNQNNNDNNVEYELIDSHSPY.....KLETIQTQKINPTNHTKTL 735

Scoring table:  
BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 9793542 seqs, 4134689005 residues  
Total number of hits satisfying chosen parameters: 19587084

Minimum DB seq length: 0  
Maximum DB seq length: 200000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL=frame+g2n.model -DEV=x1h  
-O=/cg2n\_1/USPTD\_SPOOL/US10783417/runat\_20012006\_095011\_23921/app\_query.fasta\_1.903  
-DB=Published\_Applications\_NA\_Main -QFMT=fastap -SUFFIX=g2n.rnpbm  
-MINMATCH=0.1 -LOOPEL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1  
-MATRIX=Dlosum62 -TRANS=human40.cdi -LIST=45 -DOCALLIGN=200 -THR\_SCORE=pct  
-THR\_MAX=100 -THR\_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORR=ext  
-HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000  
-USER=US10783417@cgn\_1\_1 1026 @runat\_20012006\_095011\_23921 -NCPU=6 -ICPU=3  
-NO MMAP -NEG SCORES=0 -WAIT -DSPBLOCK=100 -DEV TIMEOUT=120 -WARN TIMEOUT=30  
-THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7 -YGAPOP=10  
-YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications\_NA\_Main:\*

1: /cg2n\_6/ptodata/1/pubpna/US07\_PUBSCOMB.seq:\*  
2: /cg2n\_6/ptodata/1/pubpna/US08\_PUBSCOMB.seq:\*  
3: /cg2n\_6/ptodata/1/pubpna/US09a\_PUBSCOMB.seq:\*  
4: /cg2n\_6/ptodata/1/pubpna/US09b\_PUBSCOMB.seq:\*  
5: /cg2n\_6/ptodata/1/pubpna/US10a\_PUBSCOMB.seq:\*  
6: /cg2n\_6/ptodata/1/pubpna/US10b\_PUBSCOMB.seq:\*  
7: /cg2n\_6/ptodata/1/pubpna/US10c\_PUBSCOMB.seq:\*  
8: /cg2n\_6/ptodata/1/pubpna/US10d\_PUBSCOMB.seq:\*  
9: /cg2n\_6/ptodata/1/pubpna/US10e\_PUBSCOMB.seq:\*  
10: /cg2n\_6/ptodata/1/pubpna/US11\_PUBSCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3869	100.0	2208	8	US-10-783-417-1
2	3314	85.7	2235	8	US-10-782-570-1
3	3104.5	80.2	2085	8	US-10-782-570-3
4	849	21.9	4896	3	US-09-756-526A-3
5	849	21.9	4896	6	US-10-345-020-3
6	849	21.9	4896	6	US-10-342-821-3
7	759	19.6	3684	9	US-10-929-754-2

8	705	18.2	4391	9	US-10-929-754-4	Sequence 4, Appl1
9	703	18.2	3633	5	US-10-032-717-3	Sequence 3, Appl1
10	703	18.2	3633	6	US-10-414-637-3	Sequence 3, Appl1
11	703	18.2	3633	7	US-10-606-320-3	Sequence 3, Appl1
12	703	18.2	3633	8	US-10-746-914-3	Sequence 3, Appl1
13	703	18.2	6613	5	US-10-032-717-28	Sequence 28, Appl1
14	703	18.2	6613	6	US-10-414-637-28	Sequence 28, Appl1
15	703	18.2	6613	7	US-10-606-320-18	Sequence 18, Appl1
16	703	18.2	6613	8	US-10-746-914-18	Sequence 18, Appl1
17	689.5	17.8	4188	5	US-10-120-544A-5	Sequence 5, Appl1
18	689.5	17.8	4188	10	US-11-091-654-5	Sequence 5, Appl1
19	678.5	17.5	2003	5	US-10-032-717-7	Sequence 7, Appl1
20	678.5	17.5	2003	6	US-10-414-637-7	Sequence 7, Appl1
21	678.5	17.5	2022	5	US-10-032-717-17	Sequence 17, Appl1
22	678.5	17.5	2022	6	US-10-414-637-17	Sequence 17, Appl1
23	678.5	17.5	2022	7	US-10-606-320-13	Sequence 13, Appl1
24	678.5	17.5	2022	8	US-10-746-914-13	Sequence 13, Appl1
25	675	17.4	3621	5	US-10-032-717-1	Sequence 1, Appl1
26	675	17.4	3621	6	US-10-414-637-1	Sequence 1, Appl1
27	675	17.4	3621	7	US-10-606-320-1	Sequence 1, Appl1
28	675	17.4	3621	8	US-10-746-914-1	Sequence 1, Appl1
29	675	17.4	3621	10	US-11-021-115-5	Sequence 5, Appl1
30	675	17.4	4874	5	US-10-032-717-27	Sequence 27, Appl1
31	675	17.4	4874	6	US-10-414-637-27	Sequence 27, Appl1
32	675	17.4	4874	7	US-10-606-320-17	Sequence 17, Appl1
33	675	17.4	4874	8	US-10-746-914-17	Sequence 17, Appl1
34	669	17.3	4359	5	US-10-120-544A-3	Sequence 3, Appl1
35	669	17.3	4359	10	US-11-091-654-3	Sequence 3, Appl1
36	668	17.3	2088	10	US-11-018-615-5	Sequence 5, Appl1
37	663	17.1	1959	7	US-10-614-076-37	Sequence 37, Appl1
38	663	17.1	3258	10	US-11-018-615-28	Sequence 28, Appl1
39	663	17.1	3510	10	US-11-018-615-24	Sequence 24, Appl1
40	660	17.1	1959	5	US-10-232-665-3	Sequence 3, Appl1
41	658	17.0	1959	7	US-10-614-076-27	Sequence 27, Appl1
42	658	17.0	1959	7	US-10-614-076-35	Sequence 35, Appl1
43	657.5	17.0	2025	7	US-10-606-320-73	Sequence 73, Appl1
44	657.5	17.0	2025	7	US-10-606-320-79	Sequence 79, Appl1
45	657.5	17.0	2025	8	US-10-746-914-73	Sequence 73, Appl1

#### ALIGNMENTS

RESULT 1  
US-10-783-417-1  
; Sequence 1, Application US/10783417  
; Publication No. US20040216186A1  
; GENERAL INFORMATION:  
; APPLICANT: Carozzi, Nadine  
; APPLICANT: Hargiss, Tracy  
; APPLICANT: Kozziel, Michael G.  
; APPLICANT: Duck, Brian  
; APPLICANT: Carr, Brian  
; TITLE OF INVENTION: AXMI-006, A Delta-Endotoxin Gene and  
; FILE REFERENCE: 045600/274146  
; CURRENT APPLICATION NUMBER: US/10/783,417  
; CURRENT FILING DATE: 2004-02-20  
; PRIOR APPLICATION NUMBER: 60/448,806  
; PRIOR FILING DATE: 2003-02-20  
; NUMBER OF SEQ ID NOS: 15  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 1  
; LENGTH: 2208  
; TYPE: DNA  
; ORGANISM: Bacillus thuringiensis  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)...(2208)  
US-10-783-417-1  
Alignment Scores: 0  
Pred. No.: 3869.00  
Length: 2208  
Matches: 735

Percent Similarity: 100.0% Conservative: 0  
 Best Local Similarity: 100.0% Mismatches: 0  
 Query Match: 100.0% Indels: 0  
 DB: 8 Gaps: 0

US-10-783-417-2 (1-735) x US-10-783-417-1 (1-2208)

QY 1 MetAaNgInAaenAaSpaNaNgInUTrgIuIleIleAaSPeRHeIeThSeRProTyr 20  
 Db 1 ATGAATCAAAATAAAGCATATTAACGAATATGAAATTAATGATTCGATACCTCAT 60  
 QY 21 PhePAAaNaRgaNSerAaSpaNSerAaRgTyRProTyRThraAaSPProaNgInPro 40  
 Db 61 TTTCGAAACAGAAACAGTAATATTCAGATACCTTACACAAATAATCCAATCAACA 120  
 QY 41 LeuGInAaNTThraAaNTTyRgIuTrPLeuAaMeTcYsGInGlyAaNTThrGInTyRgIy 60  
 Db 121 TTACAAAACACAAATTAACAAAGAGTGCTCAATATGTGCAGGGAATACAAATATGCT 180  
 QY 61 AaPaenPheGInUThraPheAlaSeRaLaSPThrIleAlaIaValSeRaIaGlyThrIle 80  
 Db 181 GATATATTCGAGACATTTGCTAGTCTGATACAAATTTGCGAGTTAGTGAAGTACTAT 240  
 QY 81 ValSerGlyThrLeuLeuAlaGlyIleGlyIleuThrSerIleSerGlyProIleGly 100  
 Db 241 GTATCCGGTACTCTGTAGCCGGTATAGTGGCTCACTTCTATATCCGGAACCGATAGA 300  
 QY 101 IleIleGlyAlaIleIleSerPheGlyThrLeuIleThraValPheTrProAlaGly 120  
 Db 301 ATATATAGGTGCTATATATATATCTTTGTGACCTATATCACTGTCTTTGGCCCGGGA 360  
 QY 121 GIuGInAaPLeYThraValTrpThrGInPheIleYsMeGlyGluIlePheValaSPThr 140  
 Db 361 GAACAAGACAAACAGTATGACACAAATTTATAATGGAGAAATTTTGTGATACA 420  
 QY 141 ProLeuThrGInUSeRtleYsGInleuYsLeuGInThrLeuGInGlyPheArgGInIle 160  
 Db 421 CGGTATACAGAAAGACATATAACAGCTAAAGTTACAACTTTAGAGAAATTTAGACAAATA 480  
 QY 161 LeuGInSeRtyRAsnThraIleuAaSPaSPTrPAaRgLyLeuYsArgLeuGInAlaPro 180  
 Db 481 TTACAAAGCTATATATACACATTTAGATGATTTGGAATAATTAAGACTACACACTGCT 540  
 QY 181 GlyLeuProProSeRaIaLeuGInGInAlaIaLeuThrLeuYsIleArgPheIle 200  
 Db 541 GGATTAACCAACCATCATACAGCATTTACAAACAGCTGCTGACCTTAATAATAGATTAG 600  
 QY 201 AaenValHlaaenAaSPheIleArgGluIleProGlyPheGInleuGInUThrTyRlySthr 220  
 Db 601 AATGTTCACAAATGATTTATTCAGAAATATCCTGGTTCCTCAACTTGAATAAAGCG 660  
 QY 221 LeuLeuLeuProIleTyRAlaGInAlaIaAaenPheIleAaenLeuLeuGInGInIly 240  
 Db 661 CTATTAACCTATTTATGCGCAAGCTGCTAATTTTCATTTAATTTATTAACACAGCT 720  
 QY 241 AlaGInleuAlaAaSPGInUTrPaNaIlaaSPleHlaSProSeRGIInIleGInProaNaIla 260  
 Db 721 GGTGAATGGCTGATGAATGAATGCAATGATATACCTTCACAAATGAACCTATAGCT 780  
 QY 261 GlyThrSerAaSPaSPTyRtyRlySleuLeuYsGluAaenIleProLyStySerAaNTyr 280  
 Db 781 GGAACATACGATGACTATATTAACCTTTAAAGAAATATATCTTAATATATGTAACAT 840  
 QY 281 CysAlaAaenThTyRArgThrGlyLeuYsAaenLeuAaSPGInProaNaMeTyRTrp 300  
 Db 841 TGTGCAATACCTATATAGAACAGACTAAATAATCTTAGAGAGAACCAATATGAAATGG 900  
 QY 301 SerIlePheAaNaSPtyRArgTyRmeTThrIleThraValleuAaSPThrIleSeRGIIn 320  
 Db 901 AGTATATTTAATGATGATGAGATATATGACCATTAATGATTTAGATACCTCTCA 960  
 QY 321 PheSerLeuTyRAspIleYsArgTyRArgaSPSerIleGlyIleIleGluValIleGly 340

Db 961 TTTCTTATATATGATATATAAAGATATAGACATTCATATAGAGCAATATGAATGAAGCC 1020  
 QY 341 IleYsAaNgInleuThraRgGluIleTyRThrThrGInIleAaSPheAaSPaRgLeuPro 360  
 Db 1021 ATTAAGATATGAACCTCACAGAGAAATTTATACACTGAATAATTTTGAATGCTTCTCT 1080  
 QY 361 GInleuArgValGInProaNaLeuAlaThreGluTyRAsnleuThraRgAlaSeRphe 380  
 Db 1081 CAACCTAGAGTTCAACCAACTAGCTAGATGAGATTAATTTAAACAGTGAAGTTT 1140  
 QY 381 LysleuPheSerPheleuGInPheIlePheTyRThrGInAaenThraSPheGlyAaen 400  
 Db 1141 AAATATATTTATATTTATGAACAATTTATTTATATACGAAATAAATTTTCGGGAAAT 1200  
 QY 401 ArgLeuValGlyIleSeRaNaRgAaSPaIaProThTyRSeRaenThIleThrGInUThr 420  
 Db 1201 CGTTAGTGGTATTTCTATACGTATGACACTTATATACATATCTATATACGAAACT 1260  
 QY 421 LeuTyRgIuArgThrGlySeRProThrThrlySthrIleArgProPheGInUSeRtyR 440  
 Db 1261 TTATATGAGAAAGAACAGTTCAACCCACAAACAAATAAGACCATTTGAATCTTAT 1320  
 QY 441 LysValSerIleValThraSPaRgGInSeRProProValaSeRProIleGInProHlaSphe 460  
 Db 1321 AAAGTTCAATGTATGAATGATAGACAAATCACCTCTCTGTTCCCTTATCAACCAACTTT 1380  
 QY 461 IleIleAaNgInIleGluLeuTyRleuAaNgIlySeRaAaenThrLeuYsTySer 480  
 Db 1381 ATATATATCAAAATTTGAACCTTATTTATAATGGCTATCTTAACACACACTCAAAATTTCA 1440  
 QY 481 AlaGlySerleuSeRaNTyRGIaAaenThrPhePheGInPheProaRgLyLeYs 500  
 Db 1441 GCAGAGGGCTTTATATCAATTAATCAAAAACAACTTTTTCATTTCTTAGAATAAAAA 1500  
 QY 501 AaSPaNaenleuValIleAaSPProGlyYcYsSeRProaNaPheAaNTyRSeRaIle 520  
 Db 1501 GACTGAATCTAGATTAATGATCCAGTGTTCACCAACTTTAATACCTATATGATATAT 1560  
 QY 521 LeuSerHlaSPeSerleuPheThTyRSeRtyRValIleGlyLeuGInleuGInIleleu 540  
 Db 1561 TTATCCCATTTTTCATATTTACTTATTCCTATGATGATGGAATTAACACTCAAAATATTA 1620  
 QY 541 AaPThrGlyValleuGlyTyRThrHlaSeSerValaAaPaRgTyRAsnAlaIleSeRaSP 560  
 Db 1621 GATACAGGTATATGATGATGACACACAGTGTGTTGATATATAATGCAATATCAAGAT 1680  
 QY 561 LysIleIleThrMeTtleProAlaIleYsGlyAaenleuAaSPThraSPeRtyRVal 580  
 Db 1681 AAAATTAATTAACATGATCCAGCAATCAAAAGTAACAATCTTGATTAACAACTTAAGGTA 1740  
 QY 581 IleGluGlyProGlyHlaThrGlyYsAaenleuValTyRleuGInSeRGIInGlyArgLeu 600  
 Db 1741 ATTGAAGACCTGTGATATACAGAGAAACTTGGTTTATTTTACAAAGTCAAGGGCTTTA 1800  
 QY 601 GIuIleThrCySgInUThraSPaNSerThrGInSeRtyRpheIleArgLeuArgTyRAla 620  
 Db 1801 GAAATTAATGATGAACCTCAATTCACACAAATCTTATTTCAATTAAGACTTCGATATGCT 1860  
 QY 621 ThrAaNgIyAlaGlyAaenThrleuProaNaIleSeRleuThrIleProGlyValIleGly 640  
 Db 1861 ACAAAATGTGTGAAATAATCTTCTATATATCTTCAATATCCAGAGGATTAATAGA 1920  
 QY 641 IleProProGInArgLeuAaNaSPThraSPeRGIyThraNTyRAsnaSPleuGInTyR 660  
 Db 1921 ATACCACTCAACGACTCAACACACTTTTCTGTGACAAATTAATATTAATTTACAAATAC 1980  
 QY 661 GlyaSPheGlyTyRPhelInPheProSeRThraValThreuProleuAaNaRgAaNTle 680  
 Db 1981 GGAAGTTTGGGTATTTCCATTTTCCAAATTCAGTAACATTAACCTTTAAATCGAAACATA 2040  
 QY 681 ProPheIlePheAaNaRgAlaAaSPaYsAaenSeRtleuIleIleAaSPlySleGlu 700  
 Db 2041 CCATTTATATTTAATCGGAGATGTATCAAAATTTAATCAATTTGATTAATAATTTGAA 2100

QY 701 PheileProileThSerSerMeChigInaSmargLulysGlnlyLeugInlyLeugInlyThrIle 720  
DB 2101 TTTATACCAATTACTCTCTCTATGACCAAAATAGGAAAAACAAAATTTAGAACTATC 2160  
QY 721 GlnThrLySIIeasnthrPhePheThraenHlsthLyThLeu 735  
DB 2161 CAAACAAAATAATATACATTTTTCACAAATCATACAAAAACACTT 2205

RESULT 2  
US-10-782-570-1  
; Sequence 1, Application US/10782570  
; Publication No. US20040210965A1  
; GENERAL INFORMATION:  
; APPLICANT: Carozzi, Nadine  
; APPLICANT: Hargies, Tracy  
; APPLICANT: Koziel, Michael G.  
; APPLICANT: Duck, Nicholas B.  
; APPLICANT: Carr, Brian  
; TITLE OF INVENTION: AXMI-007, A Delta-Endotoxin Gene and  
; FILE REFERENCE: 045600/274144  
; CURRENT APPLICATION NUMBER: US/10/782,570  
; CURRENT FILING DATE: 2004-02-19  
; PRIOR APPLICATION NUMBER: 60/448,812  
; PRIOR FILING DATE: 2003-02-20  
; NUMBER OF SEQ ID NOS: 17  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 1  
; LENGTH: 2235  
; TYPE: DNA  
; ORGANISM: Bacillus thuringiensis  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)...(2235)  
US-10-782-570-1

Alignment Scores:  
Pred. No.: 0 Length: 2235  
Score: 3314.00 Matches: 638  
Percent Similarity: 91.3% Conservative: 32  
Best Local Similarity: 86.9% Mismatches: 60  
Query Match: 85.7% Indels: 4  
DB: 8 Gaps: 3

US-10-783-417-2 (1-735) x US-10-782-570-1 (1-2235)

QY 1 MetAaenGlnaSmAaSmAaSmAaenGlnyrgIuileileaspSerHlsthSerProtyr 20  
DB 1 GGTATTCACAAATATAT---AATATGAAATATGAGATTATGATTCACAAAGAAATTTATCTTAT 57  
QY 21 PheProaSmAaSmAaSmAaSmAaSerAaGlyrProtyrThraSmAaSmProaenGlnPro 40  
DB 58 CCTTCTAACAGAAATATGATCATCTTACATACCTTACACAAATATATCAATCAACCA 117  
QY 41 LeuGlnaenThraSmThyrlsGlnrPleuAaenMeCyasGlnGlyaSmnthrgInlyrGly 60  
DB 118 TTTACAAAACAAATTTACAAAGAGGCTCAATATGTGCAAGGAAATACAAATATGGT 177  
QY 61 AaPaenPheGlnThraPheAlaSerAlaSerThrIleAlaAlaValaSerAlaGlyThrIle 80  
DB 178 GATTAATTTGACACATTTGCTAGTGTGATACATATTCGCGATTAAGTTCAGAGTACTATT 237  
QY 81 ValSerGlyThrLeuLeuAlaGlyIleGlyGlyLeuThSerIleSerGlyProIleGly 100  
DB 238 GATATCCGATCTCTGTATCCGATAGGCTGAGCTTCACTTATATCCGACCGATAGGA 297  
QY 101 IleIleGlyAlaIleIleIleSerPheGlyThrLeuIleThraValPheTrpProIleGly 120  
DB 298 ATTAATAGGCTATATATATATCTTTTGTATCCCTATATACATGCTTTTGGCCCGGGA 357  
QY 121 GlnGlnaSmPlyeThraValTrpThrgInPheIlelySmecGlyGlnIlePheValaSerThr 140

DB 358 GAACAGACAAAAACGATATGACACAAATTTATTAATAATGGAGAAATTTTGTATAC 417  
QY 141 ProLeuThrgLuseryllysGlnleuThLeugInlyGlyPheargGlnIle 160  
DB 418 CCGTTTACAGAAAGCATTAACAGCTTAAGTTTCAAACTTTTGGAGATTTTACAAATA 477  
QY 161 LeuGlnSerTyraSmnthraLeuAaSpAaPTPaGlySleuLyAargLeuGlnAlaPro 180  
DB 478 TTAACAAGCTATTAATACAGCATTTAGATGATTCGAGAAATTAATAAGATCAAGCTCT 537  
QY 181 GlyLeuProProSerSerAlaLeuGlnGlnAlaAlaLeuThLeuLySIIeargPheGln 200  
DB 538 GGATTAACCAACCATCATCAGCATTTACACAAAGCTGCTTGAATCTTAATAATGATTTGG 597  
QY 201 AaenValHlSaSmAaPheIleargGlnIleProGlyPheGlnLeugInlyThrTyrlsThr 220  
DB 598 AATGTTCCAAATGATTTTATTCAGAAATACCTGCTTCCAACTTGAACCTTAATAAAG 657  
QY 221 LeuLeuLeuProIleTyraAlaGlnAlaAlaAaenPheHlSaLeuAaSmLeuGlnGlnGly 240  
DB 658 CTATTACTACTATTTATGCGCAGCTGCTATTTTCAATTTAATTTATTAACAAGCT 717  
QY 241 AlaGlnLeuAlaAaSpGlnTrpAaAlaAaPleHlSaProSerGlnIleGlnProaenAla 260  
DB 718 GCTGAATGGCTGATGAATGGAATCAGATATATATCTTCCAAATTCACCTATATGCT 777  
QY 261 GlyThraSerAaPaSpDyThyrlsLeuLeuLySIIeProLySerAaThy 280  
DB 778 GGAACATCAGATCATATTAATTAACCTTTAAAGAAATAATACCTTAATAATATGTAATCT 837  
QY 281 CysAlaAaenThyTyraGThrglyLeuLySaenLeuAaAaSpGlnProaenMeLySTP 300  
DB 838 TGTGCAATACTATTAAGAAAGACTAATAATAAATCTTGAAGAACTTAATGAGATG 897  
QY 301 SerIlePheAaenAaPtyrAaGlyrGlyrMeThrIleThraValLeuAaPThrIleSerGln 320  
DB 898 AGTATATTAATGAATTAATGAGATATATGATTAATGATTAATGATTAATGATTAATGATTA 957  
QY 321 PheSerLeuTyraSpIlelySaGlyrAaAaSerIleGlyGlyIleGlnVallySgLy 340  
DB 958 TTTCTTTTATGATTAATAAGATTAACAAAGATTAATAATGA---AGAAATAGCTGAC 1011  
QY 341 IleLySaenGlnLeuThraGlyIleTyrlsThrgInlyleAaSmPheAaPrgLeuPro 360  
DB 1012 ATTAATACTGAACCTTAACAGAAATTTTACATGAAATTAATTTTACACCTCTACT 1071  
QY 361 GlnLeuAaGlyAlaProaenLeuAlaThraMeGlnyryaSmLeuThraAlaSerPhe 380  
DB 1072 TACCTTGAATTAACCAACCAATCTCGTAAATGAATTAATTAACAGCTTACAGGCTT 1131  
QY 381 LysLeuPheSerPheLeuGlnIlePheTyrlsThrgInlyAaenThraSmPheGlyAaen 400  
DB 1132 AGATTAATTTTCAATTTTGAATGAATCTTAATTTTACAAAAAAGAAACGATCGGAAT 1191  
QY 401 ArgLeuValGlyIleSerAaSmAaAaPleProThyrySerAaSmnthrgInlyrGly 420  
DB 1192 CGTTTATGTTGATTCGAAATGATTAATGATTAATGATTAATGATTAATGATTAATGATTA 1251  
QY 421 LeuTyrglyGlnAaThrglySerProThrThyryThrgInlyAaPrgProaenGlnSerTy 440  
DB 1252 ATATATGAGAAAGAAAGAGTCCACCAACAAATACTTAATATCAATTTGAATCTTAT 1311  
QY 441 LysValSerIleValThraPrgGlnSerProProValSerProIle---GlnProHlSa 459  
DB 1312 AAAGTTCAATTTGATGATGACAAAGTAACTCTTACTTCCCTTTTCTTAACATATAC 1371  
QY 460 PheIleIleAaenGlnIleGlnLeuTyryrLeuAaenGlySerSerAaSmnthrLeuLySTP 479  
DB 1372 TTTACAAATTAATCAATTTGAATCTTATTAATTAATTAATCAACCTTAATTAATTAATCA 1431  
QY 480 SerAlaGlyGlySerLeuSerAaenTyrglnAaenThrPhePheGlnPheProaenGly 499  
DB 1432 TCAGCTGGGGGGAATTAATTAATGATTAATAAAAAACAATGATTTTCAATTTCTCTGTA 1491

```

QY 500 LysaapCyasmluvalilleapProglyCySerProasnpheasnaantYserHis 519
Db 1492 AAAGACTGTAACCAATTATTAATCAATTTTACCAAGCTATTAATATAGTCAAT 1551
QY 520 IleleuserHisPheSerleupherThyrSerYrValilleglyLeuglnleuIntle 539
Db 1552 ATTATATCCAGATTTTCTTATTAATTAATTCCTTAATAAATGGATTTAGCCGTAAATATA 1611
QY 540 LeuaSerThrGlyValleuglyYrThrHisSerSerValaAspArgTyraAsnAlaIleSer 559
Db 1612 TTATATACAGGTGATTCAGATTCAGACACACTAGCTTAATAGAAATATCAATATCA 1671
QY 560 AsplysIleIleThrMetIleProAlaIleYsgIyaasmluvalleapThrAsnSerIys 579
Db 1672 GATAAATAATATACATGATCCACCAATCAAAAGGTAAACAGCTTGATCAAACTCAAG 1731
QY 580 ValilleuglyProglyYrHisThrGlyYasmluvalYrleuglnSerGlnIyArg 599
Db 1732 GTAAATTGAAGACCTGGTCATACAGAGGAACTGGTTATTTACAAAGTCMAAGGGCGT 1791
QY 600 LeugluIleThrCyegluThrProasnerThrgInserYrPheIleArgleuArgTyr 619
Db 1792 TTAGAGATTACATGATGAATCTCTTAATTTACCAATCTTATTAATAGACTTGATAC 1851
QY 620 AlaThrAsnGlyAlaGlyAsnThrleupProasniIleSerleuThrIleProglyValille 639
Db 1852 GCTACAAATGGGTGCGAAATCTCTCTTAATATATCTTACCAATACAGAGATTAATA 1911
QY 640 GlyIleProProGlnIArgleuAsnAsnThrPheSerGlyYrThrAsnThrasnleugln 659
Db 1912 GGAATACCACTGCAACGACTCAACAACTTTTCTGACAAATTAATAATTTACAA 1971
QY 660 TyrGlyAspPheGlyYrPheGlnPheProSerThrValIleThrleuProleuAsnArgAsn 679
Db 1972 TACGAGATTTGGGTATTTCCAAATTTCCAGTACAGTACCTTAAATGGAATC 2031
QY 680 IleProPheIlePheAsnArgAlaAspValSerAsnSerIleleuIleIleAspIysIle 699
Db 2032 ATACCATTTATATTAATCGTCAGATGATCAATTCATTTATATCATGATTAATTT 2091
QY 700 GluPheIleProIleThrSerSerMetHISGlnAsnArgIulYglnIleleuglnThr 719
Db 2092 GAATTTAATCAATTAATCTCTCTGTGCGCAAAATTAAGAAAAAATTAAGAAACT 2151
QY 720 IleGlnThrIysIleAsnThrPhePheThrasnHisThrIys 733
Db 2152 ATCCAAACAAAAATAATACATTTTTCACAAATCATACAAA 2193

RESULT 3
US-10-782-570-3
; Sequence 3, Application US/10782570
; Publication No. US20040210965A1
; GENERAL INFORMATION:
; APPLICANT: Carozzi, Nadine
; APPLICANT: Hargiss, Tracy
; APPLICANT: Koziele, Michael G.
; APPLICANT: Duck, Nicholas B.
; APPLICANT: Carr, Brian
; TITLE OF INVENTION: AXMI-007, A Delta-Endotoxin Gene and
; TITLE OF INVENTION: Methods for Its Use
; FILE REFERENCE: 045600/274144
; CURRENT APPLICATION NUMBER: US/10/782,570
; CURRENT FILING DATE: 2004-02-19
; PRIOR APPLICATION NUMBER: 60/448,812
; PRIOR FILING DATE: 2003-02-20
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 2085
; TYPE: DNA
; ORGANISM: Bacillus thuringiensis
; FEATURE:

```

```

; NAME/KEY: CDS
; LOCATION: (1) ... (2085)
US-10-782-570-3

Alignment Scores:
Pred. No.: 0
Score: 3104.50
Percent Similarity: 91.94
Best Local Similarity: 87.64
Query Match: 80.24
DB: 8 Gaps: 2

US-10-783-417-2 (1-735) x US-10-782-570-3 (1-2085)

QY 52 MetCyGlnGlyAsnThrGlnIyrcIyAspAsnPhelGlnPheAlaSerAlaAspThr 71
Db 1 ATGTGTCAAGGAATATACCAATATATGATATATTTTCAGACACTTTGCTAGTCTGATACA 60
QY 72 IleAlaIleValSerAlaGlyThrIleValSerGlyYrThrleuValleGlyIleGly 91
Db 61 ATTCGTGAGTTAGTGCAGTACTATGTATCCGGTACTCTGTAGCCGGTATAGTGGG 120
QY 92 LeuThrSerIleSerGlyProIleGlyIleIleGlyValaIleIleIleSerPheGlyThr 111
Db 121 CTCACCTCTATATCCGACCATAGAAATATAGGTGTATTAATAATATCTTTGGTACC 180
QY 112 LeuIleThrValPheTrpProAlaGlyGlnIleAspIysThrValIleThrGlnPheIle 131
Db 181 CTATACACTGCTCTTTGGCCCCGGGAGACACAAACAAACAGTATGACACAAATTTAT 240
QY 132 LysMetGlyGluIlePheValaAspThrProleuThrgInserIleYsgIleuIleuIysleu 151
Db 241 AAAATGGGAGAAATTTTGTGATACCCGTTAACAGAAACATATAACAGCTAAAGTTTA 300
QY 152 GlnThrleuglyGlyPheArgGlnIleleuGlnIserYrAsnThrAlaIleuAspAspTrp 171
Db 301 CAACCTTTAGAGGATTTTAGCAAAATATTAACAAGCTTAATATACGATTAAGATATTGG 360
QY 172 ArgIysleuIysArgleuGlnAlaProGlyleuProProSerSerAlaIleuGlnIle 191
Db 361 AGAATAATTAATAAGACTCAAGCTCTCGATTTACCAACATCATCGATTCACAAACACT 420
QY 192 AlaIleuThrleuIysIleArgPheGlnAsnValHisAsnAspPheIleArgGluIlePro 211
Db 421 GCCTTGACTCTTAATAATTCAGATTTGAGAAATGTTCCACATGATTTTATTCGAGAAATACCT 480
QY 212 GlyPheGlnleuGlnThrYrIleThrleuIleuProIleYrAlaGlnAlaIleAsn 231
Db 481 GGTTCCAACTTGAAACTTATTAACCGCTATTTACTTATTTATGCGCAAGCTGCTAAT 540
QY 232 PheHisleuAsnleuIleuGlnGlnIyAlaGlnleuAlaAspGluTrpAsnAlaAspIle 251
Db 541 TTTTCATTTAATTTATTAACAACAGTCTGATTTGGCTGATGAATGGAATGACATATA 600
QY 252 HisProSerGlnIleGluProAsnAlaGlyThrSerAspAspYrYrIlyleuIleuIys 271
Db 601 CATCCTTCACAAATTTGAACCTTAAGCTGGAACATCAGATGACTATTATTAACCTTTAAAA 660
QY 272 GluAsnIleProIySerYrSerAsnYrCyalaAsnThrYrArgThrglyleuIysAsn 291
Db 661 GAAATATATACCTTAATATATAGTATTTGCTCAATATCTTAATAGAGAGACTTAATATA 720
QY 292 LeuArgAspGluProAsnMetIysTrpSerIlePheAsnAspYrYrArgArgTyrMetThr 311
Db 721 CTTGGAACGAACTTAATATAGATGAGTATTTATTAATGATTTATCGAAGATATATGACT 780
QY 312 IleThrValleuAspThrIleSerGlnPheSerleuYrAspIleYrArgYrArgAsp 331
Db 781 ATTAAGTATTAAGTATGATCGCTCAATTTCTTTTATGATATTAAGATATACAAAGAT 840
QY 332 SerIleGlyGlyIleGluValIyGlyIleYsAsnGlnleuThrArgGluIleYrThr 351
Db 841 TCAATAGAG-----AGATAGGTGGCATTAACCTGAACCTTAACAGAAATTTATACA 894

```

```

QY 352 ThrGluIleAsnPheAspArgLeuProGlnLeuArgValGlnProAsnLeuAlaThrMet 371
DB 895 ACTGAAATTAATTTTGAACCGCTCTACTTACCTTGAATTCACCCATCTCCGTATAATG 954
QY 372 GluTyrAsnLeuThrArgAlaSerPheIleuPheSerPheLeuGlnGlnPheIlePhe 391
DB 955 GAATATTAATTTTAAACAGCTTACAGGGCTTAAATATTTTCAATTTTATGATGAATTAATTT 1014
QY 392 TyrThrGlnAsnThrAsnPheGlyAsnAlaGlyIleSerAsnArgAspAlaPro 411
DB 1015 TATACAAAAAATGAAACGTACGCGAATCGTTAGTTGATTCGCAATCGTATGATGATCT 1074
QY 412 ThrTyrSerAsnThrIleThrGlnThrLeuTyrGlyAlaArgThrGlySerProThrThr 431
DB 1075 ACTTATGTCACACAGAACTGAATTAATTAATGAGAAAGAAAGAGTCCACCCACACAA 1134
QY 432 IyethrIleArgProPheGlnSerTyrIleValSerIleValThrAspArgIleSerPro 451
DB 1135 AAAACCTTAATGCAATTTGAATCCATATAAGTTTCAATTTGATGATGATGATGATGAT 1194
QY 452 ProValSerProIle---GlnProHisPheIleIleAsnGlnIleGluLeuTyrLeuAsn 470
DB 1195 CCTACTTCCCTTTCTTAACATATCTTACAAATTAATCAATTAATCAATTAATTAAT 1254
QY 471 GlySerSerAsnAsnThrLeuIleTyrSerAlaGlyIleSerLeuSerAsnTyrGlnAsn 490
DB 1255 AATTCCCTGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1314
QY 491 ThrThrPhePheGlnPheProArgIleGlyAspCysAsnLeuValIleAspProGlyCys 510
DB 1315 ACAACTGATTTTCAATTTCTCTGTAATAAAGACCTGAACCAATTAATTAATTAATTAAT 1374
QY 511 SerProAsnPheAsnAsnTyrSerHisIleLeuSerHisPheSerLeuPheThrTyrSer 530
DB 1375 TTACCAAGCTAATAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1434
QY 531 TyrValIleGlyLeuGlnLeuGlnIleLeuAspThrGlnValLeuGlyTyrThrHisSer 550
DB 1435 TATAAATTTGATTTAGCGCTAATATTAATTAATTAATTAATTAATTAATTAATTAATTA 1494
QY 551 SerValAspArgTyrAsnAlaIleSerAspIleIleIleThrMetIleProAlaIleIle 570
DB 1495 AGCTGTAATAGAAATTAATCAATATCAATTAATTAATTAATTAATTAATTAATTAATTA 1554
QY 571 GlyAsnAsnLeuAspThrAsnSerIleValIleGlyIleProGlnIleIleThrGlyIleAsn 590
DB 1555 GGTAACAGCTTGAATCAAACTCTAAGGTAATTAAGGACCTGTCATACAGAGAGAAAC 1614
QY 591 LeuValIleLeuGlnIleSerGlnIleGlyArgLeuGlnIleThrCysGlnThrProAsnSerThr 610
DB 1615 TTGGTTTATTTTCAAAAGTCAAGGCGCTTGAAGATTAACATGTAAGAACTCTTAATTTTACA 1674
QY 611 GlnSerTyrPheIleArgLeuArgTyrAlaThrAsnGlyAlaGlyAsnThrIleuProAsn 630
DB 1675 CAATCTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1734
QY 631 IleSerLeuThrIleProGlyValIleGlyIleProProGlnIleArgLeuAsnAsnThrPhe 650
DB 1735 AATATCTTCAATTAACAGAGATTAATGAATTAACCACTCAACAGACTCAACAACTTTT 1794
QY 651 SerGlyThrAsnTyrAsnAsnLeuGlnIleTyrIleAspPheGlyTyrPheGlnPheProSer 670
DB 1795 TCTGGTACAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1854
QY 671 ThrValThrLeuProLeuAsnArgAsnIleProPheIlePheAsnArgAlaAspValSer 690
DB 1855 AAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1914
QY 691 AsnSerIleLeuIleIleAspIleIleGluPheIleProIleThrSerSerMetHisGln 710
DB 1915 AATTCAATTTTATCAATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1974

```

```

QY 711 AsnArgIleGlyGlnIleLeuGlnIleThrIleGlnThrIleAsnThrPhePheThrAsn 730
DB 1975 AATAGAGAAAAACAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2034
QY 731 HisThrIle 733
DB 2035 CATACAAAA 2043
RESULT 4
US-09-756-526A-3
; Sequence 3, Application US/09756526A
; Patent No. US20020038005A1
; GENERAL INFORMATION:
; APPLICANT: Jana, Wojciechowska
; APPLICANT: Evgeny, Lewitchin
; APPLICANT: Lucmila, Revina
; APPLICANT: Igor, Zalunin
; APPLICANT: Galina, Chestukhina
; TITLE OF INVENTION: NOVEL DELTA-ENDOTOXINS AND NUCLEIC ACID SEQUENCES CODING THEREFOR
; FILE REFERENCE: S-30913A
; CURRENT APPLICATION NUMBER: US/09/756,526A
; CURRENT FILING DATE: 2001-01-08
; PRIOR APPLICATION NUMBER: US 60/175,158
; PRIOR FILING DATE: 2000-01-07
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 3
; LENGTH: 4896
; TYPE: DNA
; ORGANISM: Bacillus thuringiensis
; FEATURE:
; NAME/KEY: source
; LOCATION: (1)..(4896)
; OTHER INFORMATION: Bacillus thuringiensis subsp. finitimus strain VKPM B-1161 (taxon
; OTHER INFORMATION: :29337
; NAME/KEY: CDS
; LOCATION: (1129)..(4458)
; OTHER INFORMATION: product: Cry28Aa1 delta-endotoxin
; PUBLICATION INFORMATION:
; AUTHORS: Wojciechowska, et al.
; TITLE: Two novel delta-endotoxin gene families cry26 and cry28 from Bacillus
; JOURNAL: FEBS Lett.
; VOLUME: 453
; ISSUE: 12
; PAGES: 46-48
; DATE: 1999-06-18
; DATABASE ENTRY DATE: ---
US-09-756-526A-3
Alignment Scores:
Pred. No.: 1.54e-79 Length: 4896
Score: 849.00 Matches: 229
Percent Similarity: 48.8% Conservative: 106
Best Local Similarity: 33.3% Mismatches: 248
Query Match: 21.9% Indels: 104
DB: 3 Gaps: 25
US-10-783-417-2 (1-735) x US-09-756-526A-3 (1-4896)
QY 68 SerAlaAspThrIleAlaValSerAlaGlyThrIleValSerGlyThrIleuAsnAla 87
DB 1210 AGCAGTGAATACGTCGCTGATGAAGCGAGGATTTGATGCGTACTATCTGACAA 1269
QY 88 GlyIleGlyIleuThrSerIleSerGlyProIleGlyIleIleGlyAlaIleIleIle 107
DB 1270 -----GCCTTGCATCAATTTGTTAATCA-----GGTGTGCTACTATA 1308
QY 108 SerPheGlyThrIleuIleThrValaPheThrProAlaGlyIleGlnAsp---LysThrVal 126
DB 1309 TCATTTGGAACCTTGCTCCCGTTCTTTGGCTGATCCAGAGAAATTCACAAAAAATTT 1368
QY 127 TrpThrGlnPheIleIleuMetGlyIleIlePheValaAspThrProLeuThrGlnSerIle 146

```

```

Db      1369 TGGTACAAATTATGAAACACGAGAGAGACCTTTTAAATCAAAACAATTCTACAGCTTA 1428
Qy      147  LysGlnLeuLysLeuGlnThrLeuGlnLysPheArgGlnIleLeuGlnSerThrAsnThr 166
Db      1429 AAGAAATAGCATTCAGTTCATCTTAATAGTCTTTAAAGATGATTAACGTACTATATAAGA 1488
Qy      167  AlaLeuAspAspTyrPargLysLeuLysArgLeuGlnAlaProGlyLeuProSerSer 186
Db      1489 GCATTATATGATTGGAAGAGA-----AATCCAGTGA 1521
Qy      187  AlaLeuGlnGlnAlaAlaLeuThrLeuLysIleArgPheGlnAsnValHisAsnAspHe 206
Db      1522 -----AATACCTGCAGATTGGTATCCACAGAAATTTGAAAGCTCATTTCAATTT 1572
Qy      207  IleArgGlnIleProGlyPheGlnLeuGlnThrTyrLysThrLeuLeuProIleTyr 226
Db      1573 GTACCAATATGCCCACTCCACTCCACGTATGACACTTATTTAAGTTGCTAT 1632
Qy      227  AlaGlnAlaAlaAsnPheHisLeuAsnLeuLeuGlnGlnGlyAlaGlnLeuAlaAspGlu 246
Db      1633 ACAGAAAGCTGCAAAATTATCATTTGATTATTAACAAGGTGTAACAATTGCGGATCA 1692
Qy      247  TrpAsnAlaAspIleHisProSerGlnIleGlnProAsnAlaGlyThrSerAspTyr 266
Db      1693 TGGAAATGCAGATCAACACATTCACCAATGTGAAGTCAATCAGTACT-----TAT 1743
Qy      267  TyrLysLeuLeuLysGlnLysAsnIleProLysTyrSerAsnTyrCysAlaAsnThrTyrArg 286
Db      1744 TATGACAGCTATTTGGTATATATTTGAAAAGTATATTAATTATTTGACCAAGCATATCAT 1803
Qy      287  ThrGlyLeuLysAsnLeuArgAspGlnProAsnMetLysTyrSerIlePheAsnAspTyr 306
Db      1804 AAGAGATTGAATCACTTAAAGATCAGAAAAATATCATGAGGTGATTAATACATAT 1863
Qy      307  ArgArgTyrMetThrIleThrValIleuAspThrIleSerGlnPheSerLeuTyrAspIle 326
Db      1864 CGTCAAGAAATGACCTTATTTGATTTGATCTTGTCGCAACTTTCTTTTATATATTA 1923
Qy      327  LysArgTyrArgAspSerIleGlyGlyIleGlnValLysGlyIleLysAsnGlnLeuThr 346
Db      1924 CGTCGTTT-----CCAGAGAGGTGAGAACTGAACTGAAATTTAACA 1959
Qy      347  ArgGlnIleTyrThrThrGlnIleAsnPheAspArgLeuProGlnLeuArgValGlnPro 366
Db      1960 AGAGAGGTTTATCAAGTTTATGATCATTTAACAAGCCACA----- 2001
Qy      367  AsnLeuAlaThrMetGlnTyrAsnLeuThrArgAlaSerPheLysLeuPheSerPheLeu 386
Db      2002 -----GGGCTATTATCTTGGCTG 2019
Qy      387  GlnGlnPheIlePheTyrThrGlnAsnThrAsnPheGlyAsnArgLeuValGlyIleSer 406
Db      2020 TCAGATATTGATTTACACGAGAGGTGCGCAAGAGCGATTTATTCAGGTATTT--- 2076
Qy      407  AsnArgAspAlaProThrTyrSerAsnThr-----IleThrGlnThrLeuTyrGly 423
Db      2077 ---CGAGAGCTTAATATATATATATCGTATCAATTTTACAGATGAAAAAATATTATGCT 2133
Qy      424  GlnArgThrGly---SerProThrThrLysThrIleArgProPheGlnSerTyrLysVal 442
Db      2134 AATACAAATAGATTAAAGTAAAGCGCTCATATTATTAACAGGGGAA---TTTATGACT 2190
Qy      443  SerIleValThrAspArgIleSerProProValSerProIleGlnProHisPheIleIle 462
Db      2191 CACTTAAGCATTAACCGTCTTTTCAAAACAATAGCTGGTATAAATTAAGTTATACGTTTA 2250
Qy      463  AsnGlnIleGlnLeuTyrLeuAsnGlySerSerAsnAsnThrLeuLysTyrSerAlaGly 482
Db      2251 ATTCAAAAAATCGTATTCACAACCTTTTAAAAACGATAAT-----GAATATCAAAAAAAT 2304
Qy      483  GlySerLeuSerAsnTyr-----GlnAsnThrThrPhePheGlnPheProArgLys 499

```

```

Db      2305 TTTAATGTAATTAATCAAAATGAACCTCAGAAACTACA----- 2343
Qy      500  LysAspCysAsnLeuValIleAspProGlyCysSerProAsnPheAsnAsnTyrSerHis 519
Db      2344 -----AACTATCCAAATGATTAATGATGCTTCA---AACAGCAAAAAATTCAAACAT 2391
Qy      520  IleLeuSerHisPheSerLeuPheThrTyrSerTyrValIleGlyLeuGlnLeuGlnIle 539
Db      2392 AATTATCTCATTTTCCATTATCATCCAC-----AAGTTGAGCTT 2433
Qy      540  LeuAsp-----ThrGlyValLeuGlnIleThrThrHisSerSerValAspArgTyr 555
Db      2434 GCTAGATATTTTCACTCATATTTGCAATGAGTGGACACAAATAGTGTAACTCCCA 2493
Qy      556  AsnAlaIleSerAspLysIleIleThrMetIleProAlaIleLysGlyAsnAsnLeuAsp 575
Db      2494 AATTATATATGAAAGGTGATGATGACAAATCCCATTTGGTAAAGCTTAGGAAGTT--- 2550
Qy      576  ThrAsnSerLysValIleGlnGlyProGlyHisThrGlyLysAsnLeuValTyrLeuGln 595
Db      2551 ACTAAACAATTCAGTTATTAAGAGACACAGTTTACAGGTGAGATTTAATGAACTTCGT 2610
Qy      596  SerGlnGlyArgLeuGlnIleThrCysGlnThrProAsnSerThrGlnSerTyrPheIle 615
Db      2611 GAT-----AATGTTCTATTAATTAATGTAAGCT---AGTCTTTAAAAAAATAGCTATA 2661
Qy      616  ArgLeuArgTyrAlaThrAsnGlyAlaGlyAsnThrLeuProAsnIleSerLeuThrIle 635
Db      2662 AGTCTATTTTATGCTGCAAAATAGCAATTAAGTATCAATATCAATATGATGATTCGGA 2721
Qy      636  ProGlyValIleGlyIleProProGlnArgLeuAsnThrPheSerGlyThrAsnTyr 655
Db      2722 GCAGAGAGTTCTA-----TTGCACCTTACTCTTTCTAGAAAAGGGAAC 2763
Qy      656  AsnAsn-----LeuGlnTyrClyLysPheGlyTyrPheGlnPheProSer 670
Db      2764 AATAATTTTACAAATTCAAACCTTAACATTAAGATTTTCAATATCATACACTTTAAGT 2823
Qy      671  ThrValThrLeuProLeuAsnArgAsnIleProPheIlePheAsnArgAlaAsp----- 688
Db      2824 GATATGAAATTAACCGAAGAGAAAGAAATTCATATCCATTTGAAGCGAGAGATTAAT 2883
Qy      689  ValSerAsnSerIleLeuIleIleAspLysIleGlnPheIleProIleThrSerSerMet 708
Db      2884 GAGGAGGAGTGAATCTTTTATGATTAATAGAGTCAAAACCTATGATGAATAATAT 2943
Qy      709  HisGlnAsnArgGlnLysGlnLysLeuGlnThrIleGlnThrLysIleAsnThrPhePhe 728
Db      2944 -----ACTAATGAATGAATTTAGAGAAAGCAAGCAAGCATGTAATTAATTT 2994
Qy      729  ThrAsnHisThrLysThrLeu 735
Db      2995 AATAACGCAACAAACGCTTTG 3015

RESULT 5
US-10-345-020-3
Sequence 3, Application US/10345020
Publication No. US20030150018A1
GENERAL INFORMATION:
APPLICANT: Jana, Wojciechowska
APPLICANT: Evgeny, Lewitin
APPLICANT: Ludmila, Revina
APPLICANT: Igor, Zalunin
APPLICANT: Galina, Chesukhina
TITLE OF INVENTION: NOVEL DELTA-ENDOTOXINS AND NUCLEIC ACID SEQUENCES CODING THEREFOR
FILE REFERENCE: S-30913B
CURRENT APPLICATION NUMBER: US/10/345,020
CURRENT FILING DATE: 2003-01-15
PRIOR APPLICATION NUMBER: US 60/175,158
PRIOR FILING DATE: 2000-01-07
NUMBER OF SEQ ID NOS: 5
SOFTWARE: PatentIn version 3.0
SEQ ID NO 3

```

```

/ LENGTH: 4896
/ TYPE: DNA
/ ORGANISM: Bacillus thuringiensis
/ FEATURE:
/   NAME/KEY: source
/   LOCATION: (1)..(4896)
/   OTHER INFORMATION: Bacillus thuringiensis supsp. finitimus strain VKPM B-1161 (taxon
/   OTHER INFORMATION: 29337
/   FEATURE:
/   NAME/KEY: CDS
/   LOCATION: (1129)..(4458)
/   OTHER INFORMATION: product: Cry28Aa1 delta-endotoxin
/   PUBLICATION INFORMATION:
/   AUTHORS: Wojciechowska, et al.
/   TITLE: Two novel delta-endotoxin gene families cry26 and cry28 from Bacillus
/   TITLE: thuringiensis sp. finitimus
/   JOURNAL: FEBS Lett.
/   VOLUME: 453
/   ISSUE: 12
/   PAGES: 46-48
/   DATE: 1999-06-18
/   US-10-345-020-3
/   US-10-783-417-2 (1-735) x US-10-345-020-3 (1-4896)

```

## Alignment Scores:

```

Pred. No.: 1,54e-79 Length: 4896
Score: 849.00 Matches: 229
Percent Similarity: 48.8% Conservative: 106
Best Local Similarity: 33.3% Mismatches: 248
Query Match: 21.9% Indels: 104
DB: Gaps: 25

```

US-10-783-417-2 (1-735) x US-10-345-020-3 (1-4896)

```

QY 68 SerAlaSPThrIleAlaAlaValSerAlaGlyThrIleValSerGlyThrLeuLeuAla 87
DB 1210 AGCAGTATACAGTCTGCTAGTAAGCGAGGATGAGTGTGGTACTATACTGACA 1269
QY 88 GlyIleGlyLeuThrSerIleSerGlyProIleGlyIleIleGlyAlaIleIle 107
DB 1270 -----GCCTTGCAATCATTTGTAATCCA-----GCTGTGACTATA 1308
QY 108 SerPheGlyThrLeuIleThrValPheTrpProAlaGlyGluGlnAsp---LysThrVal 126
DB 1309 TCATTGGAACTTGGCTCCGCTCTTGGCTGATCCAGAGAGATCCAAAAAATT 1368
QY 127 TrpThrGlnPheIleLysMetGlyGluIlePheValAspThrProLeuThrGluSerIle 146
DB 1369 TGGTCACATTTATGAAACACGAGAGACCTTTTAATCAACAATTTCTACAGCTGA 1428
QY 147 LysGlnLeuLeuGlnIleThrLeuGluIlePheArgGlnIleLeuGlnSerIleThr 166
DB 1429 AAGAAATGATGATGCTCATCTAAATGTTTTAAAGATGTATTAAGTACTATGAGAA 1488
QY 167 AlaLeuAspAspTrpArgLysLeuLysArgLeuGlnAlaProGlyLeuProProSerSer 186
DB 1489 GCATTTAATGATTTGAAAGAA-----AATCCAACTGCA 1521
QY 187 AlaLeuGlnGlnAlaAlaLeuThrLeuLysIleArgPheGluAsnValHisAsnAspHe 206
DB 1522 -----AATACGCGCAGATTTGATATCAGACAGATTTGAAAGCTCATTTCAATTTT 1572
QY 207 IleArgGluIleProGlyPheGlnLeuGluIleThrIleThrLeuLeuLeuProIleIle 226
DB 1573 GTAAGCAATATGCAAACTCCAACTCCACGCTATGACACATTTATTTAAGTTGCTAT 1632
QY 227 AlaGlnAlaAlaAsnPheHisLeuAsnLeuLeuGlnGlnIleValaGluLeuAlaAspGlu 246
DB 1633 ACAAGAGCTGCANAATTACATTGATTTATTTATTCATCAAGGTTACAAATTCGGGATCA 1692
QY 247 TrpAsnAlaAspIleHisProSerGlnIleGluProAsnAlaGlyThrSerAspArgIle 266
DB 1693 TGGAAATGACGATCAACCAATTCACCAATGTTGAAGTCAATCAGGTACT-----TAT 1743

```

```

QY 267 TyrIleLeuLeuLysGluAsnIleProLysIleSerAsnIleCysAlaAsnThrIleArg 286
DB 1744 TATGACGAGCTATGTTATATATGAAAAGTATATTAATTTATGACCAACCAATCCAT 1803
QY 287 ThrGlyLeuLysAsnLeuArgAspGluProAsnMetLysTrpSerIlePheAsnAspIle 306
DB 1804 AAGAGATGAATCACCTTAAAGAAATCAGAAAAAATCACATGAGATGCTTATACATAT 1863
QY 307 ArgArgIleTrpMetThrIleThrValLeuAspThrIleSerGlnPheSerLeuIleAspIle 326
DB 1864 GCTCGAAGAAATGACCTTATATGATATGATCTTGCGCACTTTCCTTTATATATATA 1923
QY 327 LysArgIleArgAspSerIleGlyIleGluValLysGlyIleLysAsnGluLeuThr 346
DB 1924 CCTCCTTT-----CCAAAGAGATGAACTAGATTAATTAACA 1959
QY 347 ArgGluIleIleThrThrGluIleAsnPheAspArgLeuProGlnLeuArgValGlnPro 366
DB 1960 AGAGAGATTATACAAATTGATGATCATTTAACACGACCA----- 2001
QY 367 AsnLeuAlaThrMetGluIleThrAsnLeuThrArgAlaSerPheLysLeuPheSerPheLeu 386
DB 2002 -----GGCTATTTACTTGCTG 2019
QY 387 GluGlnPheIlePheThrThrGluAsnThrAsnPheGlyAsnArgLeuValGlyIleSer 406
DB 2020 TCAAGATATGAGTATATACAGAGAGTGGCGAAGCGATTTATTAACAGTAT 2076
QY 407 AsnArgAspAlaProThrIleSerAsnThr-----IleThrGluThrLeuIleGly 423
DB 2077 ---CGAGATCTAAATATATATCTGTAATCAATTTTAAATGATAAAATATTTATGCT 2133
QY 424 GluArgThrGly---SerProThrIleThrIleArgProPheGluSerIleVal 442
DB 2134 AATACAAATAGATTAAAGTAAAGACGCTCATTAATTTATTAACAGCGAA---TTATGACT 2190
QY 443 SerIleValThrAspArgGlnSerProProValSerProIleGlnProHisPheIleIle 462
DB 2191 CACTTAAGCATTAACCGTCTTTTCAACAAATGCTGGATTAATAAGTTATACACTTA 2250
QY 463 AsnGlnIleGluLeuLeuAsnGlySerSerAsnAsnThrLeuLysIleSerAlaGly 482
DB 2251 ATTCAAAAAATGATATTCACAACTTTTAAACGATAT-----GAATATCAAAAAAAT 2304
QY 483 GlySerLeuSerAsnIle-----GlnAsnThrIlePhePheGlnPheProArgLys 499
DB 2305 TTTAATGCAATTAATCAAAATGAACTCAAGAACTCA----- 2343
QY 500 LysAspCysAsnLeuValIleAspProGlyCysSerProAsnPheAsnAsnIleSerHis 519
DB 2344 -----AATATCTTAATATTAATGATGCTCA---AACAGCAAAAAATTCAAACAT 2391
QY 520 IleLeuSerHisPheSerLeuPheThrIleIleValIleGlyLeuGlnIle 539
DB 2392 AATTTATCTCATTTTCATTAATCATCAAC-----AAGTTAAGCTT 2433
QY 540 LeuAsp-----ThrGlyValLeuGlyIleProIleHisSerSerValAspArgIle 555
DB 2434 GCTGATATTTTCACTATATTTTCACTAGCTTGAGACACAAATAGTATACTCCCA 2493
QY 556 AsnAlaIleSerAspLysIleIleThrMetIleProAlaIleLysGlyAsnAsnLeuAsp 575
DB 2494 AATTTAATATCAAGAAAGTGTGATCAAAATCCCTGCTAAAGCTTACCAAGTT 2550
QY 576 ThrAsnSerLysValIleGluIleProGlyHisIleArgIleLysAsnLeuValIleGln 595
DB 2551 ACTTAACAATTCATTAATAAGAGACAGAGTTTACAGGAGGATTAATACAACTTCCT 2610
QY 596 SerGlnGlyArgLeuGluIleThrCysGluIleProAsnSerThrGlnSerIlePheIle 615
DB 2611 GAT-----AATGTCTTATTAATGATAAGCT---AGTTCTTTAAAAAATACCTATA 2661
QY 616 ArgLeuArgIleValIleThrAsnGlyValaGlyAsnThrLeuProAsnIleSerLeuThrIle 635

```





```

QY 407 AsnArgAspAlaProThrTyrSerAsnThr-----IleThrGluThrLeuTyrGly 423
    |||:::|||||
Db 2077 ---CGAGAGCTTAATATATTAATCTGGTAACTCAATTTTTCACATGAAAAATTTTATGGT 2133
QY 424 GluArgThrGly---SerProThrThrIleArgProPheGluSerTyrIleVal 442
    |||:::|||||
Db 2134 AATACAAATAGATTAAGTAAGAGCGCATTAACCTATTATTCAGGGGAA---TTTATGACT 2190
QY 443 SerIleValThrAspArgGlnSerProPheValSerProIleGlnProHisPheIleIle 462
    |||:::|||||
Db 2191 CACTTAAGCATTAACCGCTCTTTTCACAAATAGCGGTATTAATTAAGTTAATACAGTTTA 2250
QY 463 AsnGlnIleGluLeuTyrLeuAsnGlnSerSerAsnAsnThrLeuTyrTyrSerIleGly 482
    |||:::|||||
Db 2251 ATTCAAAAAATGCTATTCACAACTTTTAAAAACGATAT-----GAATATCAAAAAAAT 2304
QY 483 GlySerLeuSerAsnTyr-----GlnAsnThrThrPhePheGlnPheProArgIys 499
    |||:::|||||
Db 2305 TTATATGTGAATATATCAAAATGAGACCTCAAGAACTACG----- 2343
QY 500 LysAspCysAsnLeuValIleAspProGlyCysSerProAsnPheAsnAsnTyrSerHis 519
    |||:::|||||
Db 2344 -----AACTATCTTAATGATTAATGCGGTGCTCA---AACAGCCAAAAAATTCAAACAT 2391
QY 520 IleLeuSerHisPheSerLeuPheThrTyrSerTyrValIleGlyLeuGlnIleGlnIle 539
    |||:::|||||
Db 2392 AATTTATCTCATTTTCATTAATCATCCAC-----AAGTTAGAGTTT 2433
QY 540 LeuAsp-----ThrGlyValLeuGlyTyrThrHisSerSerValAspArgTyr 555
    |||:::|||||
Db 2434 GGTGAGTATTTTCACTCTATATTTGGATTAAGTTGGACACACATAGTGTAACTCCCA 2493
QY 556 AsnAlaIleSerAspIleIleIleThrMetIleProAlaIleIleGlyAsnAsnLeuAsp 575
    |||:::|||||
Db 2494 AATTTAATATCAGAAAGTGTGATACACAAATCCCATGTTAAAGCTTACCAAGTT--- 2550
QY 576 ThrAsnSerLeuValIleGluGlyProGlyHisThrIleGlyAsnLeuValTyrLeuGln 595
    |||:::|||||
Db 2551 ACTAACATTCAGTTATTAAGAGACCGAGGTTTTCAGGTGAGATTTAATCAACTTCCT 2610
QY 596 SerGlnGlyArgLeuGluIleThrCysGluThrProAsnSerThrGlnSerTyrPheIle 615
    |||:::|||||
Db 2611 GAT-----AAATGTTCTTATTAATGTAAGCT---AGTTCTTTAAAAAATACGCTATA 2661
QY 616 ArgLeuArgTyrAlaIleThrAsnGlyIleGlyAsnThrLeuProAsnIleSerLeuThrIle 635
    |||:::|||||
Db 2662 AGTCTATTTTATGCTGCAAAATAGCAATAGCTGATCAATAGACGTAGTGAATCCGGA 2721
QY 636 ProGlyValIleGlyIleProProGlnArgLeuAsnAsnThrPheSerGlyThrAsnTyr 655
    |||:::|||||
Db 2722 GCAGGAGTCTA-----TTGCACCTACCTTTTCTTGAAGAAAGGGAAC 2763
QY 656 AsnAsn-----LeuGlnTyrGlyAspPheGlyTyrPheGlnPheProSer 670
    |||:::|||||
Db 2764 AATATTTTACAACTTACAGACCTTAACATAAGATTTTCAATATCATACACTTTAGT 2823
QY 671 ThrValThrLeuProLeuAsnArgAsnIleProPheIlePheAsnArgIleAsp----- 688
    |||:::|||||
Db 2824 GATATTTGAATTAACCGAAGTAGAAGAAATTCATATTCATTGGAAGGAGATGATATAT 2883
QY 689 ValSerAsnSerIleLeuIleIleAspIleIleGluPheIleProIleThrSerSerMet 708
    |||:::|||||
Db 2884 GAGGAGGAGTGAATCTTTTAATTTAGTAATTAAGTTGAAACCTATTAAGTAAAAATAT 2943
QY 709 HisGlnAsnArgGlyIleGlyIleLeuGlnIleThrIleGlnThrIleAsnThrPhePhe 728
    |||:::|||||
Db 2944 -----ACTAATGAAATGATTTTAGGAAGGCAAGAAAGACAGTGAATGATATATT 2994
QY 729 ThrAsnHisThrIleThrLeu 735
    |||:::|||||
Db 2995 ATTAACGCAACAAACGCTTTG 3015

```

```

RESULT 7
US-10-929-754-2
; Sequence 2, Application US/10929754
; Publication No. US20050124803A1
; GENERAL INFORMATION:
; APPLICANT: ABDULLAH, MODH AMIR
; APPLICANT: DEAN, DONALD H.
; TITLE OF INVENTION: INSECTICIDAL CRYSTAL PROTEINS WITH ENHANCED TOXICITY
; FILE REFERENCE: 22727/04179
; CURRENT APPLICATION NUMBER: US/10/929,754
; PRIOR FILING DATE: 2004-08-30
; PRIOR APPLICATION NUMBER: 60/498,826
; PRIOR FILING DATE: 2003-08-29
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: Patent In Ver. 3.2
; SEQ ID NO 2
; LENGTH: 3684
; TYPE: DNA
; ORGANISM: Bacillus thuringiensis
US-10-929-754-2

```

```

Alignment Scores:
Pred. No.: 4,936-70 Length: 3684
Score: 759.00 Matches: 234
Percent Similarity: 47.4% Conservative: 122
Best Local Similarity: 31.2% Mismatches: 267
Query Match: 19.6% Indels: 128
DB: Gaps: 34

```

US-10-783-417-2 (1-735) x US-10-929-754-2 (1-3684)

```

QY 23 AsnArgAsnSerAsnAspSerArgTyrProTyrThrAsnAsnProAsnGlnProLeuGln 42
    |||:::|||||
Db 154 AATATGAATTCAGGC-----TATCCGTTAGCGAATGACTTCAAGGCTCAATGAATA 204
QY 43 AsnThrAsnTyrIleGlyIleThrLeuAsnMetCysGlnGlyAsnThrGlnTyrGlyAspAsn 62
    |||:::|||||
Db 205 AACGCAACTATTAAGATTTGCTGACCATGTCTGAAAAATTAACCAACAGATGCGCTTAAT 264
QY 63 PheGluThrPheAlaSerAlaAspThrIleAlaAlaValSerAlaGlyThrIleValSer 82
    |||:::|||||
Db 265 -----CCAGCTGCGGATTAATCTCTCAGTTAGTACCGCTTTAAAGTACCT 312
QY 83 GlyThrLeuLeuAlaGlyIleGlyIleThrIleThrSerIleSerGlyProIleGlyIleIle 102
    |||:::|||||
Db 313 GAGAGTATC-----CTTAATTTGTAACCAACCTGACGAGTACTGTC 354
QY 103 GlyAlaIleIleIleSerPheGlyThrLeuIleThrValPheTyrProAlaGly---Glu 121
    |||:::|||||
Db 355 TTAACCGTACTT-----AGCGGAGTCTCTCTATTTCTTGCCGCACTAATACTCA 405
QY 122 GlnAspIleThrValTyrThrGlnPheIleIleIleIleIleIleIleIleIleIleIle 141
    |||:::|||||
Db 406 ACGCTGAAGAAGTTTGAATGATTTTCATGACCAATACAGGAGATCTTAATGATCAAACT 465
QY 142 LeuThrGluSerIleIleGlyIleLeuIleIleIleIleIleIleIleIleIleIleIle 161
    |||:::|||||
Db 466 GTAACAGCTTATGTAAGAACAGATGCAATGCAAAATGACCGTTGTGAAGATTAATTA 525
QY 162 GlnSerTyrAsnThrAlaIleuAspAspTyrPheArgIleLeuIleIleIleIleIleIle 181
    |||:::|||||
Db 526 GATCAATATTAACATTAATTTTAACACTTGGAATAA----- 561
QY 182 LeuProProSerSerAlaIleGlnGlnAlaIleIleIleIleIleIleIleIleIleIle 201
    |||:::|||||
Db 562 GAGCTATTAATCAACGCTCTATGTAAGACAGATTAATCACT-----CAATTT---AAC 609
QY 202 ValHisAsnAspPheIleArgGluIleProGlyPhe-----GlnLeuGluThrTyrIys 219
    |||:::|||||
Db 610 TTAACCACTGCCAAACCTTCGAGAGACCGCAGTTTATTTAGCAACTTATGATGATGA 669
QY 220 ThrLeuLeuLeuProIleTyrAlaGlnAlaIleAsnPheHisIleuAsnIleuLeuGln 239
    |||:::|||||

```

```

Db 670 TTATTGATTATTCACATATACGACAGTAGCAAAATTCATTACTTTATATAGAAAT 729
Qy 240 GYALAGLUEVALAASPGLUTRPAENMLAEPILIEHISPROSEGLINLEGLUPROASN 259
Db 730 GGGCTCATTAATATGACACAGAAATGCTCT-----TTACACAGTACT 768
Qy 260 AAGLYTHRSERASPSPTYRTRYLSLEULEULYSGLUASMLPEPOLYSTRSERASN 279
Db 769 GCTGCT-----GACCACTATATATACACTATGTCAGTACATTAAGAAATATATGCA 822
Qy 280 TYRCSALAASNTHTYRTHRGTHGLYLEULYSAENLEUARGPGLUPROASMETLYS 299
Db 823 CATACACTATACATGGATATATAAGGTTAGTACTTGAATAAATCTATATAGACA 882
Qy 300 TTPSERILEPHEASAPRYARGARGTYRMECTHRIETHVALLEUASPTHRISESER 319
Db 883 TGGATTACGTTTAAATGATTATTAAGAGAGATGACTATTCAAGTATATATCTCGCT 942
Qy 330 GINPHESELEUTYRASPILLEYSAARGTYR---ARGASPERILEGLYGLYLEGLUVAL 338
Db 943 CTTTGTGCGAGTTATGATCAGTCGATACCTGCGGCAAAATAGATATACGAACATA 1002
Qy 339 LYSGLYLELYASAENGLUEUTHRARGULLEYRTHRTHTHGLULLEASNPHASPARG 358
Db 1003 TCA-----AAAACAGAAATTTACAGAGATTTATACAGCT-----1038
Qy 359 LEUPROGLINLEUARGVALGINPROASNLEUALATRMETGLUTYRASNLEUTHRARGALA 378
Db 1039 TTAGTAGAATCTCCTTCCTTAGTAATCTATAGCAGCACTGAGAGCACTTACACAGA---1095
Qy 379 SERPHELYSLEUPHESERPHLEUGLINPHEILEPHETYRTHGLUASNTHRANPHE 398
Db 1096 GATGTTCAATTTATTCACCTTGCTTAAAGAGAGTAGATTTCTGAGCC---AATCTATATAT 1152
Qy 399 GLYASNAAGLUEVALGLYLERASNAPARGAPALAPROTHTYRSEASNTHRIELNHR 418
Db 1153 CAAGATTTAAGATTTTATCTGCAATATAAATTTGGGTTTCAATATACAAATTTCTGCA 1212
Qy 419 -----GLUTHTREUTYRGLYGLUARG---THRGYSESPROTHRTHTRYSTRHILE 434
Db 1213 ATGCAGAAAGATGCAATTTATGAGAGTTCTGTTTGGTTCAAAATCTTACTATCAAAAT 1272
Qy 435 ARGPROPHLEUGLUSER---TYRILEVALSERILEVALTHRASPARGLINSEPROBOVAL 453
Db 1273 CAACTTAATTTCAATGTTATATAAATCTTATC---ACAGATACTAGTCCCC-----1323
Qy 454 SERPROILEGLINPROHISPHLEILELEANGLINLEGLUEUTYRLEUANGLYSEDER 473
Db 1324 -----TCT 1326
Qy 474 AENASNTHTREULYS-----TYRSEALAGLYSERLEUSERASANTYRGLINASN 490
Db 1327 AATCGAGTTACAAAATGATTTCTACAAAATGATGATGCTTGCCCTTATATATCA 1386
Qy 491 -----THRTHPHEPHEGLINPHEPROARGLYSLS 500
Db 1387 AATATACACCAACTCTGAGAGTTTAAAGACACACATTTTGGATTTTCAACAATAG 1446
Qy 501 AEPCYSAENLEUVALILEAPPROGLYCYSERPROASNPHASANTYRSEHISILE 520
Db 1447 AAC-----ACACCTAAT---CAACCAACTGTAAAGATTTATATCGCATAT 1488
Qy 521 LEUSERHISPHESERLEUPHERHTYRSETRYVALILEGLYLEUGLINLEU 540
Db 1489 TTAACCTAT-----ATMAAACTGATGTATTA 1515
Qy 541 ASPTRHGLY-----VALLEUGLYTRPHTHISERSEVALASPARGYRASN 556
Db 1516 GATTATACAGTAACAGGGTTTCATTGCTTGACACACATATGTTGTTACCTATATAT 1575
Qy 557 ALALIESERAPLYLEILETHMECTILEPROALILELYSGLYASNAENLEUASPTH 576
Db 1576 CAATATACACAGATGCTATCACACAGATTCGGCGGTAAATCTTAATCTTTGAATGCA 1635

```

```

Qy 577 AENSERLYSVALILEGLULYPROGLYHISTHRTGLYGLYASNLEUVALTYRILEUGLINSE 596
Db 1636 ACAGCTAATAGTATACAGAGACCTGCTCATACAGGGGGGAGTCTGATCTTACAAAC 1695
Qy 597 -----GINGLYARGLEUGLINLEUTHRCYSGIUTHRPRO-----AENSER 609
Db 1696 AATGCTACTCATACAGAGACAGATGAGATTCAAATTTAAACAGATATTTATATGATCCT 1755
Qy 610 THRGINSETRYRPHLEILARGLEUARGTYRVALATHRANGLYALAGLYASNTHTREUPRO 629
Db 1756 ACAGAAAGTTACGGATATACGATACGATGCTGCAAAATAGTCCA-----ATTGATATG 1809
Qy 630 AENILESELEUTHRIEPROGLYVALILEGLYILEPROGLINARGLEUASNAENTHR 649
Db 1810 AATGATATATATGATTTACAGAGACTTCTTACAGAGACACAGATTAAGTACAGAACTTACG 1869
Qy 650 PHESERGLYTHRASNTYR-----AENASNLEUGLINTYRGLYASPPHEGLYTRYRPH 666
Db 1870 TTTTCAGACCTAATATATATATATACCTACACATTTTAAATATAGAGAGTTTATGATACAA 1929
Qy 667 GIN---PHEPROSETRHTVALTHREUPROLEUASARGASN-----ILEPROHEILE 683
Db 1930 GATCCTTTGATGCAATTTGATCCGATGAGATTATCTTATACAACTGATTAATCTATAGCT 1989
Qy 684 PHEASNARGALASAPVAL---SERASNSERILELEULEILEAPLYLEGLUPHEILE 702
Db 1990 ATTCACACATTAATACATGACTTCAATATATACAGATTTATGACAGATCAAAATATAT 2049
Qy 703 PROILETHRSERSEMECHIEGLASNARGGLYULYSGINLYSLEUGLUTHRIEGLINTHR 722
Db 2050 CCAATCACTCATCATCTGATTTAGATGACAGACAGAAACCAAAATTTAGATCAGAACAGAA 2109
Qy 723 LYSILEASNTHTRPHETHTRASNTHTRYLS 733
Db 2110 GTTGAAATGACCTGTTTACAAATGACGCGAAA 2142

RESULT 8
US-10-929-754-4
/ Sequence 4, Application US/10929754
/ Publication No. US20050124803A1
/ GENERAL INFORMATION:
/ APPLICANT: ABDULAH, MODH AMIR
/ TITLE OF INVENTION: INSECTICIDAL CRYSTAL PROTEINS WITH ENHANCED TOXICITY
/ FILE REFERENCE: 22727/04179
/ CURRENT APPLICATION NUMBER: US/10/929,754
/ PRIOR FILING DATE: 2004-08-30
/ PRIOR APPLICATION NUMBER: 60/498,826
/ NUMBER OF SEQ ID NOS: 43
/ SOFTWARE: Patent In Ver. 3.2
/ SEQ ID NO 4
/ LENGTH: 4391
/ TYPE: DNA
/ ORGANISM: Bacillus thuringiensis
US-10-929-754-4

Alignment Scores:
Pred. No.: 4,48e-64 Length: 4391
Score: 705.00 Matches: 225
Percent Similarity: 43.4% Conserved: 119
Best Local Similarity: 28.4% Mismatches: 278
Query Match: 18.2% Indels: 171
DB: 9 Gaps: 29

US-10-783-417-2 (1-735) x US-10-929-754-4 (1-4391)
Qy 1 METLENGLINASNAENASPARASNGLUTYRGLULILEASPSERTHTHRSERPROTYR 20
Db 719 ATGCATATATTAATGGAATGGAATGAAATATATGACATATTA-----757
Qy 21 PHEPROASNARGAENSERASNAPSER-----ARGTYRPROTYR 33

```

Db 758 -----AATGCTCATCAACGATTCAACATGCTTAATACCTTATCCGAGTATCCGTTA 811  
 QY 34 ThrAsnAsnProAsnGlnProLeuGlnAsnThrAsnTyrLeuGlnTyrLeuAsnMetCys 53  
 Db 812 GCAAATCCACCAACAAAGATTAAATGCAAAATTAATTAAGATGGCTTAATGATATGT 871  
 QY 54 GlnGlyAsnThrGlnTyrGlyAsnAsnPheGlnThrPheAlaSerAlaAspThrIleAla 73  
 Db 872 GAAGGGTAT-----CATATAGAAAATCCATAGAGAACCAAGCGTTAGAGCT 916  
 QY 74 AlaValSerAlaGlyThrIleValSerGlyThrLeuLeuAlaGlyIleGlyIleuThr 93  
 Db 917 GCACCTGGTAAGGAGATTAGTATAGTATAGTATAGTATAGTATAGTATAGTATAGTAT 973  
 QY 94 SerIleSerGlyProIleGlyIleIleGlyAlaIleIleIleSerPheGlyThrLeuIle 113  
 Db 974 AATTATCTTATGATACATATGGATTG-----TTTATCCAAATATTCA 1012  
 QY 114 ThrValPheThrProAlaGlyGlnGlnAspLysThrValTyrThrGlnPheIleLysMet 133  
 Db 1013 GAGCTATCTTGGCCAGAGATGATACCCAGCAATACATTCGGCAAGATATTATGAATCAT 1072  
 QY 134 GlyGlnIlePheValAspThrProLeuThrGlnSerIleLysGlnLeuLysLeuGlnThr 153  
 Db 1073 GTAGAAGATCTTATAGCAAAAGAAATACAGTACGTTATACGAGAAATCATATGACACA 1132  
 QY 154 LeuGlnGlyPheArgGlnIleLeuGlnSerTyrAsnThrAlaLeuAspAspTyrArgLys 173  
 Db 1133 TTAGCAGATTATACAGGGTAAAGTTGATGATTATTAACAATTTGGTGAAGAAATGAAAGAC 1192  
 QY 174 LeuLysArgLeuGlnAlaIleProGlyLeuProProSerSerAlaLeuGlnGlnAlaAlaLeu 193  
 Db 1193 -----GATCCAAATCTACAGGTAAATTAAAGC----- 1219  
 QY 194 ThrLeuLysIleArgPheGlnLysValIleAsnAspPhe-----IleArgGlnIle 210  
 Db 1220 ACCTTAGTACCAAGATTATACGCTCTGATTCAGATTTTAATGTCCTTAAGACAGATT 1279  
 QY 211 -----ProGlyPheGlnLeuGlnTyrTyrLysThrLeuLeuProIle 225  
 Db 1280 AATTAATCAGGGAGATCCAGGTATATGAGTTA-----CTTTATTTGCTCTGTC 1324  
 QY 226 TyrAlaGlnAlaAlaAsnPheIleLeuAsnLeuLeuGlnGlnIleAlaGlnLeuAlaAsp 245  
 Db 1335 TATGCAAAATTCGCAATCTGATCTTATTAATTAAGGATGCTCAGATTTATGAGAGAT 1384  
 QY 246 GlnTyrAsnAlaAspIleHisProSerGlnIleGlnProAsnAlaGlyThrSerAspAsp 265  
 Db 1385 AATATG-----TGAGCGCACGAGCTAATGCTCTGATATAT 1420  
 QY 266 TyrTyrLysLeuLeuLysGlnLysAlaProLysTyrSerAsnTyrCysAlaAsnThrTyr 285  
 Db 1421 TATTCACCAATATCAATTAAGAAACAAAGAAATATACAGATATATGTTAATATGTTAT 1480  
 QY 286 ArgThrGlyLeuLysAsnLeuArgAspGlnProAsnMetLysTyrSerIlePheAsnAsp 305  
 Db 1481 AATTAAGGTTTAAATGATTTTAGAACACAGAGT-----CAATGGGTAAACTTATATCT 1534  
 QY 306 TyrArgArgTyrMetThrIleThrValLeuAspThrIleSerGlnPheSerLeuTyrAsp 325  
 Db 1535 TATCTGTAAGAAATGATCTTACTGATTAATGATTAATTAATTAATTAATTAATTAATGAC 1594  
 QY 326 IleLysArgTyrArgAspSerIleGlyIleGlnValIleGlyIleLysAsnGlnLeu 345  
 Db 1595 GCGAGTTATATCTCTACAGAA-----GTAAAAACCGAATCA 1630  
 QY 346 ThrArgGlnIleTyrThrThrGlnIleAsnPheAspArgLeuProGlnLeuArgValGln 365  
 Db 1631 ACTTAGGAAATTTATTCAGATGTTATTAATGGGAGATATATGACCTTATGACTCTTAT 1690  
 QY 366 ProAsnLeuAlaThrMetGlnTyrAsnLeuThrArgAlaSerPheLysLeuPheSerPhe 385

Db 1691 TTTTCTTTGAGAAAGCTGAATCACTTATATCAAGGACCC-----CATCTCTTCACTTGG 1747  
 QY 366 LeuGlnGlnPheIlePheTyrThrGlnLysAsnThrAsnPheGlyAsnArgLeuValGlyIle 405  
 Db 1748 CTAAAGATTCAGATTTGTATACCAATCTATTTCTTATGACATTTTATACAGGTGCT 1807  
 QY 406 SerAsnArgAspAlaProThrTyrSerAsnThrIleThrGln----- 419  
 Db 1808 CAATATAGATATCTTATATCTAATATATCTGATTAACGAGCGCTTTTAGGGGACAG 1867  
 QY 420 ---ThrLeuTyrGlyGlnArgThrGlySerProThrThrLysThrIleArgProPheGln 438  
 Db 1868 GACACAGATTAATGT-----GGGACTTCTTACCATTAATTAATTT-----CATCAAT 1915  
 QY 439 SerTyrLysValSerIleValThrAspArgGln-----SerProPro 452  
 Db 1916 TCGTATGTATATTAATTAATGACGAAATTAATGAATTAATTAATTAATTAATTAATTAAT 1975  
 QY 453 ValSerProIleGlnProHisPheIleIleAsnGlnIleGlnLeuTyrLeuAsnGlySer 472  
 Db 1976 GTAAATATATCAAAATGAATTTTCTGTACA-----GATAATAT 2017  
 QY 473 SerAsnAsnThrLeuLysTyrSerAlaGlySerLeuSerAsnTyrGlnAsnThrThr 492  
 Db 2018 TCTTCAAAAGATTAATTAATGATGTCACACAGAAAGAAATTAACCTGTTGTCGACAGAT 2077  
 QY 493 PhePheGlnPheProArgLysLysAspCysAsnLeuValIleAspProGlyCysSerPro 512  
 Db 2078 TTTGATTTTCTCACATTAATTAAGAGGAACTGAGTTAGCA----- 2116  
 QY 513 AsnPheAsnAsnTyrSerHisIleLeuSerHisPheSerLeu----- 526  
 Db 2117 AATATATATGATTAATCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2176  
 QY 527 -----PheThrTyrSerTyrValIleGlyLeuGlnLeuGlnIleLeuAspThrGly 543  
 Db 2177 CAGAAACGTCAATGTTATTCGTT----- 2200  
 QY 544 ValLeuGlyTyrPThrHisSerSerValAspArgTyrAsnAlaIleSerAspLysIleIle 563  
 Db 2201 -----GCTTTTACATAGTAGTGTATCTTAATTAATCACTTACGACGCAATTAAT 2254  
 QY 564 ThrMetIleProAlaIleLysGlyAsnLeuValIleLeuAsnLeuAspThrAsnSerLysValIleGlnGly 583  
 Db 2255 ACGCAATTCCTGTATGTAAGCTTCGAGTAAATGATCAATTCATTAATAAAAGGT 2314  
 QY 584 ProGlyHisThrGlyLysAsnLeuValTyrLeuGlnSerGlnIleArgLeuGlnIleThr 603  
 Db 2315 CCCGATTTTACGGGAGAGATTTGGTAAAGATGAGACAGATTCAGGTTTAATCTATGCT 2374  
 QY 604 CysGlnThrProAsnSerThrGlnSerTyrPheIleArgLeuArgTyrAlaThrAsn--- 622  
 Db 2375 TTTAAGCTGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2434  
 QY 623 -----GlyAlaGly-----AsnThrLeuPro 629  
 Db 2435 AGTTCTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2494  
 QY 630 AsnIleSerLeuThrIleProGlyValIleGlyIleProProGlnArgLeuAsnAsnThr 649  
 Db 2495 AATATTTCTCCACA----- 2509  
 QY 650 PheSerGlyThrAsnTyrAsnAsnLeuGlnTyrGlyAspPheGlyTyrPheGln---Phe 668  
 Db 2510 -----TATGAGCGCTTTCTTATTAATGAGTCTTT 2539  
 QY 669 ProSerThrValThr-----LeuProLeuAsnArgAsnIleProPheIlePheAsn 685  
 Db 2540 ACTATATACAGACGAAATATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2599  
 QY 686 ArgAlaAspValSerAsn---SerIleLeuIleIleAspLysIleGlnPheIlePro-I 704  
 Db 2600 AGACAGTTTGTGAAGATATACATCTCTTATTAATTAATTAATTAATTAATTAATTAATTAAT 2659

```
Qy 704 eThrSerMetHsgInAsn--ArgGluLysGlnLysLeuGluThrIleGlnThrly 723
Db 2660 AACTGATACCATTCACAGGAATATGAGGAAATATGAAATATGAAAGATCACAGGAAC 2719
Qy 723 sIleAenThrPhePheThrAenHisThrlyThrleu 735
Db 2720 ATTATATGATATTTGTTGATTAAACAAAGTACTA 2756

RESULT 9
US-10-032-717-3
/ Sequence 3, Application US/10032717
/ Publication No. US20020151709A1
/ GENERAL INFORMATION:
/ APPLICANT: Andre R. Abad
/ APPLICANT: Nicholas B. Duck
/ APPLICANT: Xiang Feng
/ APPLICANT: Ronald D. Flannagan
/ APPLICANT: Theodore W. Kahn
/ APPLICANT: Lynn E. Sims
/ TITLE OF INVENTION: Genes Encoding No. US20020151709A1e1 Proteins With
/ TITLE OF INVENTION: Pesticidal Activity Against Coleopterans
/ FILE REFERENCE: 35718/237005
/ CURRENT FILING DATE: 2001-10-23
/ PRIOR APPLICATION NUMBER: 60/242,838
/ NUMBER OF SEQ ID NOS: 48
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 3
/ LENGTH: 3633
/ TYPE: DNA
/ ORGANISM: Bacillus thuringiensis
/ FEATURES:
/ NAME/KEY: CDS
/ LOCATION: (1)...(3633)
/ NAME/KEY: misc_feature
/ LOCATION: (0)...(0)
/ OTHER INFORMATION: CRY1218-2
US-10-032-717-3

Alignment Scores:
Pred. No.: 5.34e-64 Length: 3633
Score: 703.00 Matches: 219
Percent Similarity: 44.5% Conservative: 127
Best Local Similarity: 28.2% Mismatches: 289
Query Match: 18.2% Indels: 142
Db: 5 Gaps: 32

US-10-783-417-2 (1-735) x US-10-032-717-3 (1-3633)
Qy 5 AenAaPaenAenGluTyrGluIleIleAaPserHisThrSerProTyrPheProAenArg 24
Db 10 AATATATCAAAATGAAATATGAAATATATATGATGCGACACCTTCT-----ACTTCT 57
Qy 25 AenSerAenAaPser---ArgTyrProTyrThrAenAenProAenGlnProLeuGlnAen 43
Db 58 GATCAATATGATTTCAACAGATACCTTTTGCGAATGAGCAAAATGCGCTACAAAT 117
Qy 44 ThrAenTyrLysGluTyrLeuAenMetCysGlnGlyAen---ThGlnTyrGlyAenAen 62
Db 118 ATGATTTTAAAGATTAATTTAAATGCTCGCGGAAATGCTAGGAATACCTGCTTCA 177
Qy 63 PheGluThrPheAIsaSerAlaPheThrIleAlaIaIaValSerAlaGlyThrIleValSer 82
Db 178 CCTGAAGTACTGTTAGCGGACACAGAT-----GAGGCTAAGCGCGCAATTGATATGTA 231
Qy 83 GlyThrLeuLeuAlaGlyIleGlyLeuThrSerIleSerGlyProIleGlyIleIle 102
Db 232 GGTAAATTAATACTATCGGTTAGG-----GTCCATTGTT 267
Qy 103 GlyAlaIleIleIleSerPheGlyThrLeuIleThrValPheTyrProAlaGlyGln 122
```

```
Db 268 GGGCCGATAGTAGTCTTTATATCACTATATGATATTTCTGTGCGCTTCAAGGCA--- 324
Qy 123 AaPlyThrValTyrThrGlnPheIleLysMetGlyIlePheValAaPThrProLeu 142
Db 325 ---AAGACTCAATGGGAATTTTATGGAACAAGTAAGAACTCATATTAATCAAAAAATA 381
Qy 143 ThrGluSerIleLysGlnLeuLysLeuGlnThrLeuGluGlyPheArgGlnIleLeuGln 162
Db 382 GCAGATATGACAGGAATTAAGCCCTTCGGAATATGAAGGATTAAGTAAATTAATCAAA 441
Qy 163 SerTyrAenThrAlaLeuAaPserTyrPheArgLysLeuLysAaGluGlnAlaProLysLeu 182
Db 442 TTATATCTAAGCTGCGCTTGAAGAAATGAAGA--- 474
Qy 183 ProPseSerAlaLeuGlnGlnAlaIaLeuThrLeuLysIleArgPheGluAenVal 202
Db 475 AATCCAAATGGTTCA-----AGAGCTTACGAGATGTCGGAATTCGATTAATTAATCTCG 528
Qy 203 HisAenAaPheIleArgGluIleProGlyPheGlnLeuGluThrTyrLysThrLeuLeu 222
Db 529 GATAGTTTATTTACGCAATATACATGCGATCTTTTCAGTGAACAAATTTGAAGTACCATTC 588
Qy 223 LeuProIleTyrAlaGlnAlaIaAenPheHisLeuAenLeuLeuGlnGlnIleValGlu 242
Db 589 CTACAGATATATACAGGACGACCACTTCAATTTACTGTTATTAAGAGACGCTTCAAT 648
Qy 243 LeuAlaAaPgluTyrAaPheAaPheHisProSerGlnIleGluProAaPheAlaGlyThr 262
Db 649 TTGAGAAAGAAATGGGATGG-----TCTACAAACACT 681
Qy 263 SerAaPsePyrTyrLysLeuLeuLysGluAenIleProLysTyrSerAenTyrCysAla 282
Db 682 ATTAATATACATTTATGATCGTCAAAATGAACCTTACGACGAATATTCGATCATCGTGA 741
Qy 283 AenThrTyrArgThrGlyLeuLysAenLysArgAaPgluProAaPheLysPheSerIle 302
Db 742 AAGGTATGAAACGTGTTTACAAATTTAAAGGACAGGCGCTAAACAGGCTGAC 801
Qy 303 PheAaAaPyrTyrArgTyrMetThrIleThrValLeuAaPThrIleSerGlnPheSer 322
Db 802 TATTAACCAATTCGTAAGAAATGACATGACGCTTTAGATGTTGTTGCAATATTCCTCA 861
Qy 323 LeuTyrAaPheLysArgTyrArgAaPserIleGlyIleGluValIleGlyIleLys 342
Db 862 AATATATGACACGACGATAC-----CCAATGAACAAACAGCA----- 900
Qy 343 AaenGluLeuThrArgGluIleTyrThrThrGluIleAaPheAaPheArgLeuProGlnLeu 362
Db 901 ---CAACTAACAAGGAAGTATATACA-----GATCCACTGGGCGCGGTA 942
Qy 363 ArgValGln-----ProAenLeuAlaThrMetGluTyr 373
Db 943 AACGTGCTTCAATGCTTCCGTGATGACAAAGACCTTCTTGGAGCTGATTAATATCA 1002
Qy 374 AaenLeuThrArgAlaSerPheLysLeuPheSerPheLeuGlnGlnPheIlePheTyrThr 393
Db 1003 TCCGTTATTCAGCACCC---CATGATATTGATTATTAACGAGACTCAAGATGATACA 1059
Qy 394 GluAenThrAaPheGlyAaPhe-----ArgLeuValGlyIleSerAaArg 408
Db 1060 CAATCAAGACATTTCTTCGCTCGCTATATTAACATTTGGCGTGCATCAAAATTAAGC 1119
Qy 409 AaPheAaProThrTyrSerAaPheThrIleThrGluThrLeuTyrGlyIleArgThrLys 428
Db 1120 TATCATCGAATTTTATGATATATTTATTAACAGATATGAACTATCAATCAAAATCTA 1179
Qy 429 ProThrThrLysThrIleArgProPheGluSerTyr-----Lys 441
Db 1180 CACAGCACTAGTACTT---GATTTACGAATATGATATTAACAAGCGTTATCAAAA 1236
Qy 442 ValSerIleValThrAaPheGlnSerProProValSerProIleGlnProHisPheIle 461
Db 1237 GATGCGGTCCTTGATATGTTTCTGTTATACGTATATA-----TTTTTGA 1290
```

```

QY 462 ILeasnglilegluentyrileuansglyseraserasnan-----ThrleuLys 478
   ::: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1291 ATGCCAGAAAGTCAGTTTTCATGCTAAACCAATTGAATATACCGAAGAGCTTAAAG 1350
QY 479 TyrSer-----AlaglyglyserleuseranTyrglnaenThr 491
   | : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1351 TATTAATCCGGTTCCAAAGATATATATAGCGCGG----- 1383
QY 492 ThrPhepheglnPheProarglyslyPhePcyAsnleuValileasProglyCyser 511
   | : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1384 -----ACAAAGATTCGAAATTGAATTGAATTAACCTCCAGAAACTTCA 1422
QY 512 -----ProasPheasnanTyrserrHisileuSerHisPhe-----SerleuPheThr 528
   | : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1423 GATCAACCAAAATTAATAGTCATATAGCCATAGATTATGTCATATCAAGATATTCGCCGG 1482
QY 529 TyrSerTyValileglyleuGlnleuGlnileuAspThrGlyValleuGlyTrpThr 548
   | : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1483 ACGGTTCAACTACCGGATTA-----GTACCTGTATTTCTTGAGACA 1524
QY 549 HisSerSerValasPheArgTyraenAlileSerAspyleileleThrMetileProAla 568
   | : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1525 CATCGGAGTCCGATCTTATTAATGAGATTCAATCAATTAATTAATCAATTCGAGTCCGGTC 1584
QY 569 ILeuysGlyAsnAsnleu-----AspThrAsnSerlySvalileGlu 582
   | : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1585 GTAAGAGTTTCTATATTTGGCTCCCTCTATTAACGAGAGGCGCAATATATACCGTTGATCG 1644
QY 583 GlyProGlyHisThrGlyAsnleuValTyrlleuGlnSerGlnGly----- 598
   | : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1645 GGTCCCTGATTAACAGGGGGGGGATTAATAAGTATATAAGAAATGAGATTAATATATCA 1704
QY 599 -----ArgleuGlnleuThrCysGlnThrProasSerThrGlnSerTyrlleuGly 616
   | : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1705 CATATCCGTTTAAATTTCA-----GACATTAAACAAGATATATGATATGAG 1752
QY 617 LeuArgTyrlleuThrAsnGlyAlaglyAsnThrleuProasnleSerleuThrilePro 636
   | : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1753 ATTCGGTATGCTTCC-----GCTAATTAATCTGAATTTTATTAATATCTCTTGAGAA 1806
QY 637 GlyValileGlyleuProGlnArgleuAsnAsnThrPheSerGlyThrAsnTyraen 656
   | : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1807 AACGTTAAATCT---CACGCTCAAAAACATGAAATAGAGTGAACCTTAACATATAT 1863
QY 657 AsnleuGlnTyrlleuAspPheGlyTyrlleuGlnPheProSerThr-----ValThr 673
   | : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1864 AATATTAAATTATGCACTTGGCCCTATTAATTAATTAACGACAAACGACCTTCAATTA 1923
QY 674 LeuProleuAsnArgAsnleuProPhelePheAsnArgAlaAspValSerAsnSerile 693
   | : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1924 CTAGGGGCT-----ATAATTGAACCGGAAAGACTTCTTGGAATTTGA 1965
QY 694 LeuileleAspAspyleleGlnPhelePheleProleleThrSerSerMetHisGlnAsnArgGlu 713
   | : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1966 GCTTATATATGACGATTCGAATTTATCCAGTATGATGACACATAT-----GAAGCG 2016
QY 714 TyrlleuGlnleuGlnleuThrleuGlnThrlyleuAsnThrPhePheThrAsn 730
   | : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 2017 GAACAAGATTTAGAACAGCAAGAAAGACAGTGAATGCTTGTTCGAT 2067

```

```

; TITLE OF INVENTION: Pesticidal Activity Against Coleopterans
; FILE REFERENCE: 35718/237005
; CURRENT APPLICATION NUMBER: US/10/414,637
; CURRENT FILING DATE: 2003-04-16
; PRIOR APPLICATION NUMBER: US/10/032,717
; PRIOR FILING DATE: 2001-10-23
; PRIOR APPLICATION NUMBER: 60/242,838
; PRIOR FILING DATE: 2000-10-24
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 3633
; TYPE: DNA
; ORGANISM: Bacillus thuringiensis
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(3633)
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (0)...(0)
; OTHER INFORMATION: C1Y1218-2
; US-10-414-637-3

Alignment Scores:
Pred. No.: 5,346-64 Length: 3633
Score: 703.00 Matches: 219
Percent Similarity: 44.5% Conservative: 127
Best Local Similarity: 28.2% Mismatches: 289
Query Match: 18.2% Indels: 142
DB: Gaps: 32

US-10-783-417-2 (1-735) x US-10-414-637-3 (1-3633)
QY 5 AsnAspAsnAsnGlnTyrlleuGlnleuAspSerHisThrSerProTyrlleuPheProasnArg 24
   | : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 10 AATTAATCAAAATGAAATGAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 57
QY 25 AsnSerAsnAspSer---ArgTyrlleuGlnleuGlnleuGlnleuGlnleuGlnleuGlnleu 43
   | : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 58 GTATCCAAATGATCTTCAACGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 117
QY 44 ThrAsnTyrlleuGlnleuGlnleuGlnleuGlnleuGlnleuGlnleuGlnleuGlnleuGlnleu 62
   | : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 118 ATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 177
QY 63 PheGlnThrPheAlaSerAlaSerAlaSerAlaSerAlaSerAlaSerAlaSerAlaSerAlaSer 82
   | : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 178 CCTGAAGTACTTGTAGCGGACAGAT-----GCAGCTAAGCCGCAATGATATATATGA 231
QY 83 GlyThrleuLeuAlaGlylyleuGlyleuThrSerleSerGlyProleleGlylelle 102
   | : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 232 GGTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 267
QY 103 GlyAlailelleleleSerPheGlyThrleuileThrleuValPheTrpProAlaGlyGln 122
   | : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 268 GGGCGGATATGATGCTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 324
QY 123 AspTyrlleuValTyrlleuGlnleuGlnleuGlnleuGlnleuGlnleuGlnleuGlnleuGlnleu 142
   | : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 325 ---AAGAGTCATATGGAAGATTTTATGGAACAAGTGAAGAACTCATTAATCAAAAATATGA 381
QY 143 ThrGlnSerleuGlnleuGlnleuGlnleuGlnleuGlnleuGlnleuGlnleuGlnleuGlnleu 162
   | : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 382 GCAGAAATATGCAAGAAATTAAGCGCTTTCGAAATTAAGAAATTAAGAAATTAATTAATTAATTA 441
QY 163 SerTyrlleuThrAlaAspAspAspTyrlleuGlnleuGlnleuGlnleuGlnleuGlnleuGlnleu 182
   | : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 442 TTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 474
QY 183 ProProSerSerAlaGlnleuGlnleuAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAla 202
   | : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 475 AATCCAAATGCTTCA-----AGAGCTTATCAAGATGCGCAAAATCGATTTGAATATCCCTG 528

```

OY	203	HisaaNpPheIleArgGluleProGlyPheGlnleuGluThrTyrlserThreuleu	222
Db	529	GATAGTTATTATACCAATACATGCATCTTTTCAGAGACAAATTTTGAAGTACATTC	588
OY	223	LeuProIleTyrlaGlnlaalaanPheHisleuAenleuGlnGlnIvalaGlu	242
Db	589	CTTACAGATATATACACAGGACCAACCTTCATTACTGTATTTAAAGACCGCTCAATT	648
OY	243	LeuIleaaPgluTTPasmlaasplleHisProSerGlnIleGluPProasmlaGlyThr	262
Db	649	TTTGAGAGAAAGATGGGATG-----TTTCAACCAACT	681
OY	263	SerAaPspTyrlTyrlseuleuLeuylsGluAenIleProlyeTySerAntyCyala	282
Db	682	ATTATATACTATATATGATCGTAANTGAACCTTACTGCAGATATTTGATCACTGTGA	741
OY	283	AenThrTyrlaGThGlyleuLysaenleuAargaspGluProasmeIylsTrpserIle	302
Db	742	AAAGTGATGAAACCTGGTTTACCAAAATTTAAAGCACAGCGCTTAAACAAATGGGTGCAC	801
OY	303	PheaAaPpyTyrlAargArgTyrlmetThrIleThrValleuAapThrIleSerGlnPheSer	322
Db	802	TATAACCAATTCGGTAGAAATGACACGAGGTTTAAAGATGTGGCATTAATTTCCCA	861
OY	323	LeuTyrlaapIleLysleuTyrlaargaspSerIleGlyGlyIleGluValIleGlyIleLys	342
Db	862	AATTATGACACACGACGTAC-----CCAATGAACGAAGACA-----	900
OY	343	AenGluLeuThrArgGluIleTyrlThrThrgluIleasnPheaAargLeuProGlnleu	362
Db	901	---CACTAACACAGGAAGTATATCA-----GATCCACTGGGCGCGGTA	942
OY	363	ArgValGln-----ProasnleuAlaThreGluTyrl	373
Db	943	AAACGCTTCAATTGGTCTCGTAGACAAAGCACCTTCTTCGAGTGAATGAATCA	1002
OY	374	AsnleuThrArgAlaSerPheLysleuPheSerPheleuGluGlnPheIlePheTyrlThr	393
Db	1003	TCCGTTATTCCAGCACACC---CATGTATTGATTATATACGGAGCTCAGAGTATACA	1052
OY	394	GluAenThrAsnPheGlyAsn-----ArgleuValGlyIleSerasnArg	408
Db	1060	CAATCAAGAGACATTTCTCCGCTCGCTATATATAGACATATGGGCTGATCAAAATAC	1112
OY	409	AspAlaProThrTyrlSerAsnThrIleThrgluThrleuTyrlGlyIuArgThrglySer	428
Db	1120	TATCATCGAATTTTAAAGTATATATTTAAAAACAGATGATAGCAATCAAAATCTA	1172
OY	429	ProThrThrLysThrlleArgProGheGluSerTyrl-----Lys	441
Db	1180	CACAGCACTAGTACCTTT---GATTTTACGAATATATGATATTATACAGACGTTATCAAAA	1233
OY	442	ValSerIleValThrAspArgGlnSerProProValSerProIleGlnProHisPheIle	461
Db	1237	GATGGGGGTCCTTGATATGTGTTTTCTCGTTAGCTATATA-----TTTTTTGA	1292
OY	462	IleasnGlnIleGluLeuTyrlleuasnGlySerSerAsnAsn-----ThrleuLys	478
Db	1291	ATGCCAGAAAGTCAGATTTTTCATGCTTAAACCAATTGAATAATACAGAAAGCGTTAAAG	1350
OY	479	Tyrlser-----AlaGlyGlySerLeuSerAsnTyrlolnAsnThr	491
Db	1351	TATATATCCGGTTTCCAAAGATATATATAGGGGG-----	1383
OY	492	ThrPhePheGlnPheProArgLysLysaAPCyAsnleuValIleAapProGlyCysSer	511
Db	1384	-----ACAAAGATTCGGAATTGATTTACCTCCAGAAACTTCA	1422
OY	512	-----ProasnPheaAsnAntyrlSerHisIleleuSerHisPhe---SerleuPheThr	528
Db	1423	GATCAACCAAAATTATAGTCATATAGCCATATAGTATATGTCATATACAAAGTATTCGCCGG	1482
OY	529	TyrlserTyrlValIleGlyleuGlnleuGlnIleleuAapThrGlyValleuGlyTyrlThr	548

```

Db      1483  ACGGGTTCACTACCGGATT-----GTAACCTGATTTTCTTGACA 1522
Oy      549  HisSerValAspArgTyrAsnAlaLeSerAspLysIleIleThrMetIleProAla 568
Db      1525  CATCGAGATCGCGATCTTATTAATGACAGTTCATTCAGATTAATAATTACTACAGATTCGGGTC 1584
Oy      569  IleIysGlyAsnAsnLeu-----AspThrAsnSerLysValIleGlu 582
Db      1585  GTAAAGGTTTGTGATTGTGGCTTCCTCTATTAACAGAGGGCCCAATAATTAACCGTGTATCG 1644
Oy      583  GlyProGlyHisIleThrGlyLysAsnLeuValTyrLeuGlnSerGlnGly----- 598
Db      1645  GGTCTGTGATTACAGCGGGGGGATATATAAGATTAAGAAATGAGATATTATATCA 1700
Oy      599  -----ArgLeuGlnIleThrCysGluThrProAsnSerThrGlnSerTyrPheIleArg 616
Db      1705  CATATGCGTGTAAATTTCA-----GACATTAACAAAGATATAGTATGAG 1752
Oy      617  LeuArgTyrAlaThrAsnGlyAlaGlyAsnThrLeuProAsnIleSerLeuThrIlePro 636
Db      637  GlyValIleGlyIleProProGlyArgLeuAsnAsnThrPheSerGlyThrAsnTyrAsn 656
Oy      1807  AACGTTAAATCT---CACGCTCAAAAACCTATGAAATAGAGGTGAAGCTTTAACTATTAAT 1866
Db      1864  AAATTATATATGAGACTTTGCCCCCTATTAATTTACGACACCGAACCTTTCATTACT 1922
Oy      674  LeuProLeuAsnAlaGsnIleProPheIlePheAsnAlaIleArgValSerAsnSerIle 693
Db      1924  CTAGCGGCT-----ATAATTGAAGCGGAAGACTTTCTTGAATTGA 1965
Oy      694  LeuIleIleAspLysIleGluPheIleProIleThrSerSerMetHisGlnAsnArgGlu 713
Db      1966  GCTTATATAGCCGATGGAATTTATTCCTCAATGATGAGACATAT-----GAAGCG 2013
Oy      714  LysGlnLysLeuGlnIuThrIleGlnThrLysIleAsnThrPhePheThrAsn 730
Db      2017  GAACAAAGATTAGAAAGCAGCGAAGAAAGCAAGTGAATGCTTGTTACGAAT 2067

RESULT 11
US-10-606-320-3
; Sequence 3, Application US/10606320
; Publication No. US20040091505A1
; GENERAL INFORMATION:
; APPLICANT: Andre R. Abad
; APPLICANT: Ronald D. Flanagan
; APPLICANT: Rafael Herrmann
; APPLICANT: Theodore W. Kahn
; APPLICANT: Albert L. Lu
; APPLICANT: Billy Fred McCutchen
; APPLICANT: James K. Prensall
; APPLICANT: James F.H. Wong
; APPLICANT: Cao-Guo Yu
; TITLE OF INVENTION: Genes Encoding Proteins With Pesticidal
; FILE OF INVENTION: Activity
; FILE REFERENCE: 35718/263948
; CURRENT FILING DATE: US/10/606,320
; PRIOR APPLICATION NUMBER: 60/391,786
; PRIOR FILING DATE: 2002-06-26
; PRIOR APPLICATION NUMBER: 60/460,787
; PRIOR FILING DATE: 2003-04-04
; NUMBER OF SEQ ID NOS: 134
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 3633
; TYPE: DNA
; ORGANISM: Bacillus thuringiensis
; FEATURE:

```

```

; NAME/KEY: CDS
; LOCATION: (1) ... (3633)
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (0) ... (0)
; OTHER INFORMATION: CRY1218-2
;
US-10-606-320-3

```

### Alignment Scores:

Pred. No.:	5.34e-64	Length:	3633
Score:	703.00	Matches:	219
Percent Similarity:	44.5%	Conservative:	127
Best Local Similarity:	28.2%	Mismatches:	289
Query Match:	18.2%	Indels:	142
DB:	7	Gaps:	32

US-10-783-417-2 (1-735) x US-10-606-320-3 (1-3633)

Oy		5	AshasphasnengluYrYguilelleaspserHlsmhsrseProYrPhaproAsnArg	24
Db		10	AAATATCAAAATGAAATAGATAATTATAGATGGACACTTCT-----ACTTC	57
Oy		25	AsnSerAsnaSpSer---ArGYrProYrTrHaaAenProAsnInProLeuglnAsn	43
Db		58	GTAATCCAAATGATTCTAAACAGATACCCTTTTGGAATAGAGCAAACAATAGCGCTACAAAT	117
Oy		44	ThrasnTYrlsgluTripleusennetCysgInglyAsn--ThrgInTYrlyAspsn	62
Db		118	ATGATATATTAAGAATTATTTAAAAATGTCCGGGAAAATGCTAGGAATACCCGTGTCA	177
Oy		63	PhegluThrPhehlaseRhalasprHrlealaaIvalaserAlagLYrHleValSer	82
Db		178	CTGAAGCTACTGTTAGCCGACAACT-----GCAGCTAAGGCCGCAATGATATGTA	231
Oy		83	GLYrHrleuDeulaaglYlleglYgleYrHsrlerlesergYrProlelgYlle	102
Db		232	GGTAATATTACTATCACAGTTTAGGG-----GTCCATTTGTT	267
Oy		103	GLYAlalellelleeserPheglYrHrleulleHrValPheTrPProlaglygluGln	122
Db		268	GGGCGCATGATGAGCTTTATATCTAACATTATGATATCTGTGGCTTCAGGGCAA---	324
Oy		123	AsplyrThervalYrThrgInPheHrleYsmetGLYglnIlePheValAspThrProleu	142
Db		325	---AAGGTCATATGGAGATTTTATAGCAACAAGTAAAGCAACTATATAACAAAAATA	381
Oy		143	ThrgInserIleYsgInleuYsleuGlnThrleuGlnYrPheargGlnIleleuGln	162
Db		382	GCAGATATATGACAGAAATAAAGCGCTTTCGGAATTAGAAAGATTAGTAATATACAA	441
Oy		163	SerTYrAsnthrAlaleuAspaPrTPahgYrSleuYsaYrleuGlnAlarProglYleu	182
Db		442	TATATCTTAATCTCGCTTGAAGATGGAAGAA-----	474
Oy		183	ProProSerSerAlaleuGlnAlalaIaleuthrleuYsleargPheglAsnVal	202
Db		475	AATCCAAATGGTCA-----AGAGCTTACAGAGATGCGCAATTCGATTTGAAATCTCG	528
Oy		203	HIsAsnaSpHeHleargIleuLeProglYrPheGlnleuGlnYrThrTYrSthrleu	222
Db		529	GATATGTTATTTCCGCAATATGATGCATCTTTTCCAGTACGACAAATTTTGAAGTACAAATC	588
Oy		223	LeuProleYrYrAlaGlnAlalaAnPheHlsleuAenleuLeuGlnGlnYrAlaGlu	242
Db		589	CTTACAGATATATACAGAGCGAGCAACTTTCATTACTGTATTAAGACAGCTTCAATT	648
Oy		243	LeuAlaAspGIurYPaanlaIaaprIHslsProserGlnIleGluProAsnAglyYrH	262
Db		649	TTTGGAGAAAGAAATGGGATG-----TCTACAAACACT	681
Oy		263	SerAspaPrTYrTYrlySleuleuYsGluAbenIIeProYrYrTSerAsnTYrCyala	282
Db		682	ATTATATCAATATATATGATGTCMAATGAAACTTATCGCAATATATTCGATCACTGTGTA	741

OY	283	AsnThrLyxArgThrglyLeuLysAsnLeuArgAspGluProAsnMetLysTrpSerIle	302
Db	742	AAAGTGATATGAACCTGGTTTATGCAAAATTAAGAAGCAGACCGCTAAACATGGCTGCAC	801
OY	303	PheAsnAspTyraGArgTyMetThIleThrValLeuAspThrIleSerGlnPheSer	332
Db	802	TATTAACCAATTCGGTAGAAGAAATACACTACGCGTTTTATGATGTTGTGCAATTATTC	861
OY	323	LeuTyraAspIleLysArgTyraArgAspSerIleGlyGlyIleGluValLysGlyIleLys	342
Db	862	AATTATGACACACCCACAGTAC-----CCAAATGGAAAGAAAGCA-----	900
OY	343	AsnGluLeuThrArgGluIleTyThrThrGluIleAsnPheAspArgLeuProGlnLeu	362
Db	901	---CACTAACACAGGAAGTATATCA-----GATCACTGGGCGCGGT	942
OY	363	ArgValGln-----ProAsnLeuAlaThrMetGluTy	373
Db	943	AACTGTCTTCAATGTGGTCCGTGTATGACAAAGCACCTTTCGGAGTATGAATGATCA	1002
OY	374	AsnLeuThrArgAlaSerPheLysLeuPheSerPheLeuGlnPheIlePheTyThr	393
Db	1003	TCCGTATTGACCAACC---CATGTATTGTATTATTAACGGGACTCACAGTGTATACA	1059
OY	394	GluAsnThrAsnPheGlyAsn-----ArgLeuValGlyIleSerAsnArg	408
Db	1060	CAATCAAGAAGCAATTTCTCCGCTCGGTATATAGAATTTGGCGTGTATCAAAATAGC	1119
OY	409	AspAlaProThrTySerAsnThrIleThrGluThrLeuTyrglyLysArgThrglySer	428
Db	1120	TATCATCGAGATTATTAGATTAATATTATTAACAGATTTAGTGAACATAATCAAAATCTA	1179
OY	429	ProThrThrLysThrIleArgProPheGluSerTy-----Lys	441
Db	1180	CACAGCACTAGTACCTT---GATTTACGAATATATATATTACAAAGACGTATCAAA	1236
OY	442	ValSerIleValThrAspArgGlnSerProProValSerProIleGlnProHisPheIle	461
Db	1237	GATCGCGGTCTCTGTATATATTGTTTCTCGTTATACGTATATA-----TTTTTTGGA	1290
OY	462	IleAsnGlnIleGluLeuTyryLeuAsnGlySerSerAsn-----ThrLeuLys	478
Db	1291	ATGCGAAGATCGAGTTTTCATGTGTAAACCAATGTATATATCCAGAAAGACGTTTAAG	1350
OY	479	Tyiser-----AlaGlyGlySerLeuSerAsnTyrglnAsnThr	491
Db	1351	TATTAATCCGTTTCCAAAGATATTATATGCGGG-----	1383
OY	492	ThrPhePheGlnPheProArgLysLysAspCysAsnLeuValIleAspProGlyCysSer	511
Db	1384	-----ACAAAGATTCGGAATTCAAAATTAACCTCCAGAAACTTCA	1422
OY	512	-----ProAsnPheAsnAsnTySerHisIleLeuSerHisPhe---SerLeuPheThr	528
Db	1423	GATCAACCAATTAATGAGCTCATATAGCCATAGATTAATGATATACACAAAGTATTC	1482
OY	529	TySerTyraValIleGlyLeuGlnLeuGlnIleLeuAspThrGlyValLeuGlyTyThr	548
Db	1483	ACGGGTCACTACCGGATTA-----GTACCTGATATTTCTTGAC	1524
OY	549	HisSerSerValAspArgTyraAsnAlaIleSerAspLysIleIleThrMetIleProAla	568
Db	1525	CATGGAGGTGCGGATCTTATTAATGAAGCTTACATTCAGATTAATTTACTCAGATTC	1584
OY	569	IleLysGlyAsnAsnLeu-----AspThrAsnSerLysValIleGln	582
Db	1585	GTAAGGTTCGTGATTTGGCTCCCTCATATAACAGAGGCGCAATATATACCGTTGATCG	1644
OY	583	GlyProGlyHisThrGlyLysAsnLeuValTyryLeuGlnSerGlnGly-----	598
Db	1645	GGTCTCGGATTTACAGGGGGGGGATTAATAAGTATTAAGAAATGAGATTAATATATCA	1704

QY 599 -----ArgLeuGluIleThrCysGluThrProAsnSerThrGlnSerTyrrPheIleArg 616  
Db 1705 CATATGCGGTGTTAAATTTCA-----GACATTACCAAGAAATATAGTATGAGG 1752  
QY 617 LeuArgTyrrAlaThrAsnGlyAlaGlyAsnThrLeuProAsnIleSerLeuThrIlePro 636  
Db 1753 ATTCGGATGCTTC-----GCTAATTAATACGAAATTTATATTAATCTTCTGAAGAA 1806  
QY 637 GlyAlaIleGlyIleProProGlnArgLeuAsnAsnThrPheSerGlyThrAsnTyrrAsn 656  
Db 1807 AAGCTTAATCT---CACGCTCAAAAACATGAAATAGAGCGAAGCTTAAACATATAT 1863  
QY 657 AsnLeuGlnTyrrGlyAspPheGlyTyrrPheGlnPheProSerThr-----ValThr 673  
Db 1864 AATATTAAATTAATGCGACTTTGCCCCCTTATTAATTTACCAACCAACCTTCACTTACT 1923  
QY 674 LeuProLeuAsnArgAsnIleProPheIlePheAsnArgAlaAspValSerAsnSerIle 693  
Db 1924 CTAGGGGCT-----AATATTGAACCGAAGACTTCTTGAAATTGAA 1965  
QY 694 LeuIleIleAspLysIleGluPheIleProIleThrSerSerMetIleGlnAsnArgGlu 713  
Db 1966 GCTTATATAGACGAAATGCAATTTATCCAGTAGATGACACATAT-----GAAGCG 2016  
QY 714 LysGlnLysLeuGluThrIleGlnThrLysIleAsnThrPhePheThrAsn 730  
Db 2017 GAACAAAGATTATGAAGCAGCAAGAAAGCAGTGAATGCTTGTTCGAAAT 2067

RESULT 12  
US-10-746-914-3  
/ Sequence 3, Application US/10746914  
/ Publication No. US20040210963A1  
/ GENERAL INFORMATION:  
/ APPLICANT: Albert L. Lu  
/ TITLE OF INVENTION: Genes Encoding Proteins With Pesticidal  
/ TITLE OF INVENTION: Activity  
/ FILE REFERENCE: 35718/268350  
/ CURRENT APPLICATION NUMBER: US/10/746,914  
/ PRIOR FILING DATE: 2003-12-24  
/ PRIOR APPLICATION NUMBER: 60/391,786  
/ PRIOR FILING DATE: 2002-06-26  
/ PRIOR APPLICATION NUMBER: 60/460,787  
/ PRIOR FILING DATE: 2003-04-04  
/ PRIOR APPLICATION NUMBER: 10/606,320  
/ PRIOR FILING DATE: 2003-06-25  
/ NUMBER OF SEQ ID NOS: 143  
/ SOFTWARE: FastSeq for Windows Version 4.0  
/ SEQ ID NO 3  
/ LENGTH: 3633  
/ TYPE: DNA  
/ ORGANISM: *Bacillus thuringiensis*  
/ FEATURE:  
/ NAME/KEY: CDS  
/ LOCATION: (1)...(3633)  
/ FEATURE:  
/ NAME/KEY: misc\_feature  
/ LOCATION: (0)...(0)  
/ OTHER INFORMATION: Cyt1218-2  
US-10-746-914-3

Alignment Scores:  
Pred. No.: 5,34e-64 Length: 3633  
Score: 703.00 Matches: 219  
Percent Similarity: 44.5% Conservative: 127  
Best Local Similarity: 28.2% Mismatches: 289  
Query Match: 18.2% Indels: 142  
DB: 8 Gaps: 32

US-10-783-417-2 (1-735) X US-10-746-914-3 (1-3633)  
QY 5 AaaAaPaAaAaGluTyrrGluIleIleAaPSeRHiThSeRProTyrrPheProAaArg 24  
Db 10 AATTAATCAAAATGAATATGAAATTAATTAATGACACACTTCT-----ACTTCT 57

QY 25 AsnSerAsnAspSer---ArgTyrrProTyrrThrAsnAsnProAsnGlnProLeuGlnAsn 43  
Db 58 GTATCAATATGATTTACACAGATATACCTTTGGGAATGAGCCAAACAAATGCGCTACAAAT 117  
QY 44 ThrAsnTyrrLysGluThrLeuAsnMetCysGlnGlyAsn---ThrGlnTyrrGlyAsn 62  
Db 118 ATGATTTATTAAGATTAATTAATAAATGTCGGGGAATATGCTAGGAATACCTCGTTTCA 177  
QY 63 PheGluThrPheAlaSerAlaAspThrIleAlaIleValSerAlaGlyThrIleValSer 82  
Db 178 CCGTAAGTACTTGTTAGCGGACAGAT-----GACGTAAGCGCGCAATTGATATGTA 231  
QY 83 GlyThrLeuLeuAlaGlyIleGlyIleLeuThrSerIleSerGlyProIleGlyIleIle 102  
Db 232 GGTAAATTAATCACTACAGTTTAGG-----GTCCCATTTGTT 267  
QY 103 GlyAlaIleIleIleSerPheGlyThrLeuIleThrValPheTrpProAlaGlyGluGln 122  
Db 268 GGGCCGATAGAGATCTTATTAATCTCACTTATGATATCTGTGTGCTTACAGGCA--- 324  
QY 123 AspLysThrValTrpThrGlnPheIleLysMetGlyGluIlePheValAspThrProLeu 142  
Db 325 ---AAGAGTCATAGGAGATTTTATGCAACAGTAGAAGAACTCATTAATCAAAAATA 381  
QY 143 ThrGluSerIleLysGlnLeuLysLeuGlnThrLeuGlnGlyPheArgGlnIleLeuGln 162  
Db 382 GCAGAAATATGCAAGAAATTAAGCCGCTTCGGAATTAAGAGATTAGTAATTAATACCA 441  
QY 163 SerTyrrAsnThrAlaLeuAspAspTrpArgLysLeuLysArgLeuGlnAlaProGlyLeu 182  
Db 442 TTATATCTTAACCTGCTTGAAGAAATGAAGAA----- 474  
QY 183 ProProSerSerAlaLeuGlnGlnAlaIleLeuThrLysLysIleArgPheGluAsnVal 202  
Db 475 AATCCAAATGGTTC-----AGACCTTAGCAGATGTGCAAAATCGATTTAAATCTCG 528  
QY 203 HisAsnAspPheIleArgGluIleProGlyPheGlnLeuGlnThrTyrrLysThrLeuLeu 222  
Db 529 GATAGTTATTTATGACCAATACATGCGATCTTTGACAGTAGCAAAATTTGAATGACATTC 588  
QY 223 LeuProIleTyrrAlaGlnAlaAlaAsnPheHisLeuAsnLeuLeuGlnGlnGlyAlaGlu 242  
Db 589 CTACAGATATATACACAGACCGCAACCTTCACTTATTAAGGACCGCTTCATTT 648  
QY 243 LeuAlaAspGluTrpAsnAlaAspIleHisProSerGlnIleGluProAsnAlaGlyThr 262  
Db 649 TTTGGAAGAAATGGGATG-----TTCAACACACT 681  
QY 263 SerAspAspTyrrTyrrLysLeuLeuLysGluAsnIleProLysTyrrSerAsnTyrrCysAla 282  
Db 682 AATTAATACATAATTAATGATCGCAAAATGAACCTTAGCGAATATTCGATCATGTGTA 741  
QY 283 AsnThrTyrrArgThrGlyLeuLysAsnLeuArgAspGluProAsnMetLysTrpSerIle 302  
Db 742 AAGGTGATGAAACTGGTTTAGCAAAATTAAGGACGACGAGCGCTAAACAAAGGTGAC 801  
QY 303 PheAsnAspTyrrArgArgTyrrMetThrIleThrValLeuAspThrIleSerGlnPheSer 322  
Db 802 TATTAACAAATTCGTAGAGAAATGACACTGACGGTTTGAATGTTGTTGATTAATCCCA 861  
QY 323 LeuTyrrAspIleLysArgTyrrArgAspSerIleGlyGlyIleGluValLysGlyLys 342  
Db 862 AATTAATGACACGACGATAC-----CCAATGAAACGAACCA----- 900  
QY 343 AsnGluLeuThrArgGluIleTyrrThrGluIleAsnPheAspArgLeuProGlnLeu 362  
Db 901 ---CAACTTAACAAAGGAATATATACA-----GATCCACTGGCGCGGTA 942  
QY 363 ArgValGln-----ProAsnLeuAlaThrMetGluTyrr 373  
Db 943 AAGGTGCTTCAATTGTTCTGTTATGACAAAGACCTTCTTGGAGTGAATGAATCA 1002



QY	374	AsnLeuThrArgAlaSerPheLeuPheSerPheLeuGluGlnPheIlePheThr	393
Db	1003	TCGGTATTGCACCAACC---CAGTATTGATTATATACGGCACTCAAGTGTATCA	1055
QY	394	GluAsnThrAsnPheGlyAsn-----ArgLeuValGlyIleSerAsnArg	408
Db	1060	CAATCAAGAAAGCATTTCTTCGCTCCGTATATTAAGACATTTGGGCGGTGCATCAATAAGC	1119
QY	409	AspAlaProThrTyrSerAsnThrIlePheGluThrLeuTyrGlyGluArgThrGlySer	428
Db	1120	TATCATCGGATTTTATGATATATATTTAAACAGATGTATGAACTAATCAAAATCTA	1179
QY	429	ProThrTyrThrIleArgProPheGluSerTyr-----Lys	441
Db	1180	CACAGCACTAGTACTCTT---GATTATCGAATTTATGATATTTACAGACCTTATCAAAA	1238
QY	442	ValSerIleValThrAspArgGlnSerProProValSerProIleGlnProHisPheIle	461
Db	1237	GATGGCGGCTCCCTTGATATGTTTCTTCGTATATAGTATATAT-----TTTTTGGC	1290
QY	462	IleAsnGlnIleGluLeuTyrIleuAsnGlySerSerAsnAsn-----ThrIleuLys	478
Db	1291	ATGCCAGAAAGTCAGATTTTTCATGTGTAACCAATGATATATACAGAAAGCGTTAAAG	1350
QY	479	TyrSer-----AlaGlyGlySerLeuSerAsnTyrGlnAsnThr	491
Db	1351	TATATATCCGGTTCCCAAGATATATATACGGGGG-----	1383
QY	492	ThrPhePheGlnPheProArgIleYelYsAspCysAsnLeuValIleAspProGlyCysSer	511
Db	1384	-----ACAAAGATTCGGAAATTATGAAATTAACCTCCAGAAACTTCA	1422
QY	512	-----ProAsnPheAsnAsnTyrSerHisIleLeuSerHisPhe--SerLeuPheThr	528
Db	1423	GATCAACCAATATATGATCATATATATGACATATATATGTCATATCACAAGATATCCCGCG	1482
QY	529	TyrSerTyrValIleGlyLeuGlnLeuGlnIleLeuAspThrGlyValLeuGlyTyrThr	548
Db	1483	ACGGGTTCAACTACCGGATTA-----GTACTGTATTTCTTGGAGCA	1522
QY	549	HisSerSerValAspArgTyrAsnAlaIleSerAspLysIleIleThrMetIleProAla	568
Db	1525	CATCGAGATGCCGATCTTATATATATGCAATTCATTCAGATAAAATTAATCTCAGATTCGGGTC	1588
QY	569	IleIysGlyAsnAsnLeu-----AspThrAsnSerLysValIleGlu	582
Db	1585	GTAAGAAGTTTCGATTGCTGCCCTCTATTAACAGAGAGGGCCAAATATATACGGTGTATCG	1644
QY	583	GlyProGlyHisThrGlyIleAsnLeuValTyrLeuGlnSerGlnGly-----	598
Db	1645	GGTCCCTGATTATACAGGGGGGGGAGTATTAAGAAATTAAGAAATGAGTATATATATCA	1704
QY	599	-----ArgLeuGlnIleThrCysGluThrProAsnSerThrGlnSerTyrPheIleArg	616
Db	1705	CATATGCGGTGTAAATTTCA-----GACATTAACAAAGAAATATATGATATGAGG	1755
QY	617	LeuArgTyrAlaThrAsnGlyValaGlyAsnThrLeuProAsnIleSerLeuThrIlePro	636
Db	1753	ATTGCGTATGCTCTC-----GCTAATATACTGAATTTATATATTAATCTTCTGCAAGA	1800
QY	637	GlyValIleGlyIleProProGlnArgLeuAsnAsnThrPheSerGlyThrAsnTyrAsn	656
Db	1807	AACGTTAATCT---CAGCTCTCAAAAACCTATGATTAAGGTAAGGCTTTAACAATATAT	1863
QY	657	AsnLeuGlnTyrGlyAspPheGlyTyrPheGlnPheProSerThr-----ValThr	673
Db	1864	AAATTTAATATATGCACTTTGCCCCCTATTAATTTATTAATTTACACAAACCTTTCATTA	1923
QY	674	LeuProLeuAsnArgAsnIleProPheIlePheAsnArgAlaAspValSerAsnSerIle	693
Db	1924	CTAGGGGCT-----ATATTTGAAGCGGAAGACTTCTCTTGGAATGAA	1966
QY	694	LeuIleIleAspLysIleGluPheIleProIleThrSerSerMetHisGlnAsnArgGlu	713

```

Db      1966 GCTTATATAGCCGAATCGAATTATATCCAGTAGATGAGCATAT-----GAAGCG 2016
OY      714  LysGlnLysLeuGluThrIleGlnThrLysIleAsnThrPhePheThrAsn 730
Db      2017 GAAACAAGATTAGAAAGCAGCGAAGAAAGACGTGAATCCCTGTTTACGAAT 2067

RESULT 13
US-10-032-717-28
; Sequence 28, Application US/10032717
; Publication No. US20020151709A1
; GENERAL INFORMATION:
; APPLICANT: Andre R. Abad
; APPLICANT: Nicholas B. Duck
; APPLICANT: Xiang Feng
; APPLICANT: Ronald D. Flanagan
; APPLICANT: Theodore W. Kahn
; APPLICANT: Lynn B. Sims
; TITLE OF INVENTION: Genes Encoding No. US20020151709A1el Proteins with
; TITLE OR INVENTION: Pesticidal Activity Against Coleopterans
; FILE REFERENCE: 35718/23705
; CURRENT APPLICATION NUMBER: US/10/032, 717
; CURRENT FILING DATE: 2001-10-23
; PRIOR APPLICATION NUMBER: 60/242, 838
; PRIOR FILING DATE: 2000-10-24
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 28
; LENGTH: 6613
; TYPE: DNA
; ORGANISM: Bacillus thuringiensis
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (0) .. (0)
; OTHER INFORMATION: Genomic Cry12I8-2
US-10-032-717-28

Alignment Scores:
Pred. No.:      1,47e-63      Length:      6613
Score:          703.00      Matches:      219
Percent Similarity: 44.5%      Conservative: 127
Best Local Similarity: 28.2%      Mismatches:  289
Query Match:    18.2%      Indels:      142
DB:             5          Gaps:          32

US-10-783-417-2 (1-735) x US-10-032-717-28 (1-6613)
OY      5  AsnAspAsnAsnGluTyrGluIleIleAspSerIsthrSerProTyrPheProAsnArg 24
      |||:::| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      1263 AATATATCAAAATGAAATATGAAATATATGATGCGACACCTCT-----ACTTCT 1310
OY      25 AsnSerAsnAspSer---ArgTyrProTyrThrAsnAsnProAsnGlnProLeuGlnAsn 43
      |||:::| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      1311 GTATCCCAATGATTTCTTACACAGTATACCTTTTCCGAATAGACCAACAATGCGCTACAAAT 1370
OY      44 ThrAsnTyrLysGluTyrPLeuAsnMetCysGlnGlyAsn---ThrGlnTyrGlyAspAsn 62
      ::::| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      1371 ATGATGATTATAAAGATTATTATAAATGTCGTGCGGAATGCTAGTAATACCTTGCTTCA 1430
OY      63 PheGluThrPheAlaSerIleAlaAspThrIleAlaValaValaSerIleAlaGlyThrIleValSer 82
      |||:::| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      1431 CCTAAGACTACTGTTTAGCGGACAAAGAT-----GCACGTAAAGCGCGCAATTGATATAGTA 1484
OY      83 GlyThrLeuLeuAlaGlyIleGlyGlyLeuThrSerIleSerGlyProIleGlyIleIle 102
      |||:::| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      1485 GGTAAATTACTATACGATTAGCG-----GTCCCAATTGCT 1520
OY      103 GlyAlaIleIleIleSerPheGlyThrLeuIleThrValPheTyrProAlaGlyGluGln 122
      ::::| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      1521 GGGCGGATAGAGCATCTTTATATACCAACTTATATATGTTCTGTGCGCTTCAGGCGCA-- 1577
OY      123 AspLysThrValTyrPThrGlnPheIleLysMetCylGluIlePheValaAspThrProLeu 142
      |||:::| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

Db 1578 --AAGAGCAATGGGAGATTTTATNGAACAAGTAAGAAAGCACTGATAATCAAAAAATA 163

Qy ThrGluSerIleuylGlnleuLysleuGlnThrLeuGluGlyPheArgGlnIleuGln 162

Db 1635 GCAGGATATGCAAGGAATATAAGCCGCTTCGGAATTAGAAAGGATTAGTAATATTATCAAA 169

Qy SerTYRAnThrAlaLeuAspAspTrpArgLysLeuLysArgLeuGlnAlaProLysLeu 182

Db 1695 TTATATCTAACTGCGCTTGAAAGATGGAAAGA----- 1722

Qy ProProSerSerAlaLeuGlnGlnAlaAlaLeuThrLeuLysIleArgPheGluAsnAla 202

Db 1728 AATCCAAATGGTCA-----AGAGCCTTACGAGATGGCGAAATCGAATTTGAATTCGTG 178

Qy HisAsnAspPheIleArgGluIleProGlyPheGlnLeuGluThrTYrLysThrLeuLeu 222

Db 1782 GATAGCTTATTTACCCAAATACATGCACTCTTTTCCAGTGCACAAATTTGAAGTACCATTC 184

Qy LeuProIleTYrAlaGlnAlaAlaAsnPheHisLeuAsnLeuLeuGlnGlnIleuLysIle 242

Db 1842 CTACAGATATATACACGACGACCACTTCATTTACTGTTATTAAGACCGCTTCAATT 190

Qy 243 LeuAlaAspGluTrpAsnAlaAspIleHisProSerGlnIleGluProAsnAlaGlyThr 262

Db 1902 TTGGAGAAAGATGGGATG-----TCTACAAACACT 193

Qy 263 SerAspAspTYrTYrLysLeuLeuLysGluAsnIleProLysTYrSerAsnTYrCysAla 282

Db 1935 ATTAATATCTATTATGATCGTCAATGAAAGAACTTACTGCAGAAATATTTGATCATCTGTGA 199

Qy 283 AsnThrTYrArgThrGlyLeuLysAsnLeuArgAspGluProAsnMetLysTrpSerIle 302

Db 1995 AAGTGCATGAAACGCTGTGACAAATTAATAAGGACAGCGCTAAACAATGGGTGCAGC 205

Qy 303 PheAsnAspTYrArgArgTYrMetThrIleThrValIleuAspThrIleSerGlnPheSer 322

Db 2055 TATAACCAATTCGTAAGGAATATGACATGACGCGTTTNGATGCTGTGCTATATCCCA 211

Qy 323 LeuTYrAspIleLysArgTYrArgAspSerIleGlyGlyIleGluValLysGlyIleLys 342

Db 2115 AATTATGACACACGCACTGAC-----CCAATGAAACGAAAGCA----- 215

Qy 343 AsnGluLeuThrArgGlnIleTYrThrThrGluIleAsnPheAspArgLeuProGlnLeu 362

Db 2154 --CACTAACAGGGAAGTATATCA-----GATCCACTGGCGCGGTGA 219

Qy 363 ArgValGln-----ProAsnLeuAlaThrMetGluTYr 373

Db 2196 AACGCGTTCATATGTTCCCGGTATGACAAAGACCTTCTTCGGAGTGATATACAA 225

Qy 374 AsnLeuThrArgAlaSerPheLysLeuPheSerPheLeuGlnGlnPheIlePheTYrThr 393

Db 2256 TCCGTTATTCGACCAACC--CATGTATTGTATATATACGGACCTCACAGTGTATACA 231

Qy 394 GluAsnThrAsnPheGlyAsn-----ArgLeuValGlyIleSerAsnArg 408

Db 2313 CAATTAAGAGCATTTCTTCGCTGCTATATTAAGACATGGGCTGTCATCAATAATAGC 237

Qy 409 AspAlaProThrTYrSerAsnThrIleThrGluThrLeuTYrGlyGluArgThrGlySer 428

Db 2373 TATCATCGGATTTTNGATATATTTATAAACAAGATGTATGGAACATAATCAAAATCTGA 242

Qy 429 ProThrThyLysThrIleArgProPheGluSerTYr-----Lys 441

Db 2433 CACAGCACTAGACCTT--GATTTACGAATATATGATATTTACAAAGCTATACAAA 248

Qy 442 ValSerIleValThrAspArgLysSerProProValSerProIleGlnProHisPheIle 461

Db 2490 GATGCGGCGCTCTGATATATGTTTCTCGTTATAGTATATA-----TTTTTGA 254

Qy 462 IleAsnGlnIleGluLeuTYrLeuAsnLysSerAsnAsn-----ThrLeuLys 478

Db 2544 ATGCCAGAGTCGAGCTTTTCTAGGTGTAAACCAATGGAATATATACGAAAGACGTTAAG 260

QY	479	YrSer-----	AlaGlyGlySerLeuSerSerAntyGlnAenThr	491
Db	2604	TATATATCCGGTTCCAAAGATATTATACGGGG-----		263
QY	492	ThrPhePheGlnPheProArgLysLysAspCysAsnLeuValIleAspProGlyCysSer	511	
Db	2637	-----ACAAGAGATTCCGAATTGGAATTGGAATTCCTCCAGAACTTCA	267	
QY	512	-----ProAsnPheAsnAsnTyrSerHisIleLeuSerHisPhe-----SerLeuPheThr	528	
Db	2676	GATCAACCAAAATTATAGAGTCATATAGCCCATAGATATTATGTCATATCACAAGATATCCCGCG	273	
QY	529	TyrSerTyrValIleGlyLeuGlnLeuGlnIleLeuAspThrGlyValLeuGlyTPTPr	548	
Db	2736	ACGGTTCAACTACCGGATTA-----GTACCTGATATTTCTTGAGCA	277	
QY	549	HisSerSerValAspArgTyrAsnAlaIleSerAspLysIleIleThrMetIleProAla	568	
Db	2778	CATCGAGATGCCGACATCTTATAATGCAGTTCATTCAGATAAATTAATCAAGATATCCGGTC	283	
QY	569	IleLysGlyAsnAsnLeu-----AspThrAsnSerLysValIleGlu	582	
Db	2838	GTAAGGTTTCGATTGTGGCTCCCTCTATACAGAGAGGCCAAATATACCGTTGTATCG	289	
QY	583	GlyProGlyHisThrGlyValAsnLeuValTyrLeuGlnSerGlnGly-----	598	
Db	2898	GGTCTGATTTACAGCGGGGGGAGATATATAAGATATAGAAATGAGCTAATTATATATCA	295	
QY	599	-----ArgLeuGlnIleThrCysGluThrProAsnSerThrGlnSerTyrPheIleArg	616	
Db	2958	CATATGCCGTGTAAATTTCA-----GACATTACAAAGATATATGATAGAGG	300	
QY	617	LeuArgTyrAlaThrAsnGlyValGlyAsnThrLeuProAsnIleSerLeuThrIlePro	636	
Db	3006	ATTCGGTATGCTCC-----GCTATATATCTGAATTTATATATATCTTCTGAAGAA	305	
QY	637	GlyValIleGlyIleProProGlnArgLeuAsnAsnThrPheSerGlyThrAsnTyrAsn	656	
Db	3060	AACGTTAATCT---CACGCTCAAAAACCTATGATATAGACGTGAAGCTTAACTATATAT	311	
QY	657	AsnLeuGlnTyrGlyAspPheGlyTyrPheGlnPheProSerThr-----ValThr	673	
Db	3117	AAATTTATATTATGACACTTTCGCCCTTATTAATTTAGACAAACCGAACTTTCATACT	317	
QY	674	LeuProLeuAsnArgAsnIleProPheIlePheAsnArgAlaAspValSerAsnSerIle	693	
Db	3177	CTAGGGGCT-----ATATTTGAAGCGGAAGACTTCTTGGAAATTGAA	321	
QY	694	LeuIleIleAspLysIleGluPheIleProIleThrSerSerMetHisGlnAsnArgGlu	713	
Db	3219	GCTTATATATAGACCGAATTCGATTTATCCACGTATAGATGACATAT-----GAAGCG	326	
QY	714	LysGlnLysLeuGlnIleThrIleGlnThrLysIleAsnThrPhePheThrAsn	730	
Db	3270	GAACAGATATTTAGAAGACAGACGAAGAAACAGTGAATGCCCTGTTTACGAAT	3320	
RESULT 14				
US-10-414-637-28				
Sequence 28, Application US/10414637				
Publication No. US20030177528A1				
GENERAL INFORMATION:				
APPLICANT: Andre R. Adad				
APPLICANT: Nicholas B. Duck				
APPLICANT: Xiang Feng				
APPLICANT: Ronald D. Flanagan				
APPLICANT: Theodore W. Kahn				
APPLICANT: Lynn R. Sims				
TITLE OF INVENTION: Genes Encoding No. US20030177528A1 Proteins With				
TITLE OF INVENTION: Pesticidal Activity Against Coleopterans				
FILE REFERENCE: 35716/237005				
CURRENT APPLICATION NUMBER: US/10/414.637				
CURRENT FILING DATE: 2003-04-16				

RESULT 14  
US-10-414-637-28  
Sequence 28, Application US/10414637  
Publication No. US2003017528A1  
GENERAL INFORMATION:  
APPLICANT: Andre R. Abad  
APPLICANT: Nicholas B. Duck  
APPLICANT: Xiang Feng  
APPLICANT: Ronald D. Flanagan  
APPLICANT: Theodore W. Kahn  
APPLICANT: Lynn E. Sims  
TITLE OF INVENTION: Genes Encoding No. US2003017528A1el Proteins With  
TITLE OF INVENTION: Pesticidal Activity Against Coleopteraans  
FILE REFERENCE: 35718/237005  
CURRENT APPLICATION NUMBER: US/10/414,637  
CURRENT FILING DATE: 2003-04-16

```

/ PRIOR APPLICATION NUMBER: US/10/032,717
/ PRIOR FILING DATE: 2001-10-23
/ PRIOR APPLICATION NUMBER: 60/242,838
/ PRIOR FILING DATE: 2000-10-24
/ NUMBER OF SEQ ID NOS: 48
/ SOFTWARE: PaSeq for Windows Version 4.0
/ SEQ ID NO 28
/ LENGTH: 6613
/ TYPE: DNA
/ ORGANISM: Bacillus thuringiensis
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (0)...(0)
/ OTHER INFORMATION: Genomic Cry12Ia-2
US-10-414-637-28

Alignment Scores:
Pred. No.: 1,47e-63 Length: 6613
Score: 703.00 Matches: 219
Percent Similarity: 44.5% Conservative: 127
Best Local Similarity: 28.2% Mismatches: 289
Query Match: 18.2% Indels: 142
Db: 6 Gaps: 32

US-10-783-417-2 (1-735) x US-10-414-637-28 (1-6613)

QY 5 AaAspAsnAsnGluTrpGluIleIleAspSerHisThrSerProGlyrPheProAsnArg 24
Db 1263 AATAAATCAAAATGAATATGAAATTAATGATGACACCTTCT-----ACTTCT 1310
QY 25 AaAspAsnAspSer---ArgTyrProGlyrThrAsnProAsnGlnProLeuGlnAsn 43
Db 1311 GATATCAATGATTTCTAAGATACCTCTTTGGCAATGAGCCAAATGGCGTCAAAAT 1370
QY 44 ThrAsnTyrGluGluTrpLeuAsnMetCysGlnGlyAsn---ThrGlnTyrGlyAspAsn 62
Db 1371 ATGGAATTAATAAGATATTAATAAATGCTGCGGAAATGCTGAAATACCTGGTCA 1430
QY 63 PheGluThrPheAlaSerAlaAspThrIleAlaIleValSerIleGlyThrIleValSer 82
Db 1431 CCTGAAGTACTGTTAGCGGACAAAGAT-----GCAGCTAAGGCCGCAATTGATATGTA 1484
QY 83 GlyThrLeuLeuAlaGlyIleGlyIleGlyLeuThrSerIleSerGlyProIleGlyIleIle 102
Db 1485 GGTAAATTAATCTATCAGGTTTAAGG-----GTCCATTGTT 1520
QY 103 GlyAlaIleIleIleSerPheGlyThrLeuIleThrValPheTyrProAlaGlyGln 122
Db 1521 GGCGCATATGATGCTTATTAATCACTTATGATATCTGTGGCTTCAAGGCAA--- 1577
QY 123 AspIleThrValTyrPheGlnPheIleIleValMetGlyGluIlePheValAspThrProLeu 142
Db 1578 ---AAGAGTCAAAGGAGATTTTATGAAACAAGTGAAGMACCTCAATCAAAAAATA 1634
QY 143 ThrGluSerIleGlyGlnLeuIleGlnThrLeuGlnGlyPheArgGlnIleLeuGln 162
Db 1635 GCAGAAATATGCAAGAAATTAAGCGCTTTGCGAATTGAAAGATTAAGTAATTAATTCGA 1694
QY 163 SerTyrAsnThrAlaLeuAspAspTyrPargIleLeuIleValArgLeuGlnIleProGlyLeu 182
Db 1695 TTATATCTAATCGCGCTTGAGAAATGAAAGAA----- 1727
QY 183 ProProSerSerAlaLeuGlnGlnIleAlaLeuThrLeuIleValIleArgPheGluAsnVal 202
Db 1728 AATCCAAATGCTTCA-----AGAGCCTTACGAGATGTGCAAAATGATTTGAAATCTCG 1781
QY 203 HisAsnAspPheIleArgGluIleProGlyPheGlnLeuGlnThrTyrIleThrLeuLeu 222
Db 1782 GATAGTTTATTTTCGCAATATCATGCACTTTTCGATGACAAATTTGAAGTACCAATTC 1841
QY 223 LeuProIleTyrAlaGlnIleAlaAsnPheHisLeuAsnLeuLeuGlnGlnIleAlaGln 242
Db 1842 CTTACAGTATATACACAGGCAAGCACTTCATTTACTGTATTAAGGACGCTTCAAT 1901

```

```

QY 243 LeuAlaAspGluTrpAsnAlaAspIleHisProSerGlnIleGluProAsnAlaGlyThr 262
Db 1902 TTTCGAGAAAGATGGGATGG-----TCTACAACTACT 1934
QY 263 SerAspAspTyrTyrIleLeuLeuIleValIleProIleTyrSerAsnTyrCysAla 282
Db 1935 ATTAATTAATTAATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1994
QY 283 AsnThrTyrArgThrGlyLeuLeuIleValIleAspGluProAsnMetCysTyrSerIle 302
Db 1995 AAGTGATATGAACATGCTTTTACCAAAATTAAGGACGAGCGCTAACAATGGCTGCAC 2054
QY 303 PheAsnAspTyrArgArgTyrMetThrIleThrValIleAspThrIleSerGlnPheSer 322
Db 2055 TATTAACCAATTCGATGAGAAATGACATGACGCTTTTGAATGCTTGTGATTAATTCGA 2114
QY 323 LeuTyrAspIleLeuArgTyrArgAspSerIleGlyIleGluValIleGlyIleLeu 342
Db 2115 AATTATGACACACGACGCTAC-----CCAATGAAACGAAGCA----- 2153
QY 343 AsnGluLeuThrArgGluIleTyrThrThrGluIleAsnPheAspArgLeuProGlnLeu 362
Db 2154 ---CACTTAACAAGGAGATATATACA-----GATCACTGGCGCGGCTA 2195
QY 363 ArgValGln-----ProAsnLeuAlaThrMetGluTyr 373
Db 2196 AACGTTCTTCAATTCGTTCTGCTGATGACAAAGCACTTCTTTCGAGATGATGAATCA 2255
QY 374 AsnLeuThrArgAlaSerPheLeuIleAspPheLeuGlnGlnPheIlePheTyrThr 393
Db 2256 TCGGTATTCGACCAACCC---CATGATTTGATTAATAACGGGACTCACAGTATATAC 2312
QY 394 GluAsnThrAsnPheGlyAsn-----ArgLeuValGlyIleSerAsnArg 408
Db 2313 CAATCAAGAAGCATTTCTCCCTGCTATATAGAATGATGGCTGATCAATAATAGC 2372
QY 409 AspAlaProThrTyrSerAsnThrIleThrGluThrLeuTyrGlyGluIleArgThrGlySer 428
Db 2373 TATCATCGGATTTTATGATATATATATATTAACAAGATATGAACTATCAATAATCTA 2432
QY 429 ProThrThrLeuThrIleArgProPheGlnSerTyr-----Lys 441
Db 2433 CACAGACATGATACCTT---GATTTTACGATTAATGATTAACAAGCGTTATCAAAA 2489
QY 442 ValSerIleValThrAspArgIleSerProValSerProIleGlnProHisPheIle 461
Db 2490 GATGCGGTCTCTGATATGTTGTTTCTGCTTATACGTAATATA-----TTTTTGA 2543
QY 462 IleAsnGlnIleGluLeuTyrIleAsnGlnIleSerSerAsnAsn-----ThrLeuIle 478
Db 2544 ATGCCAAGATGCAAGTTTTCATGTGAACCAATTGAATAAATCCAGAAAGCGTTAAG 2603
QY 479 TyrSer-----AlaGlyGlySerLeuSerAsnTyrGlnAsnThr 491
Db 2604 TATATCCGTTTCCAAAGATATTAATAGCGGG----- 2636
QY 492 ThrPhePheGlnPheProArgIleIleValAspCysAsnLeuValIleAspProGlyCysSer 511
Db 2637 -----ACAAGAGATTCGGAATTAATGATTAATCACTCCAGAAATCTTCA 2675
QY 512 -----ProAsnPheAsnAsnTyrSerHisIleLeuSerHisPhe---SerLeuPheThr 528
Db 2676 GATCAACCAAAATTAATGATATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2735
QY 529 TyrSerTyrValIleGlyLeuGlnIleLeuAspThrGlyValLeuGlyTyrPhe 548
Db 2736 ACGGTTCAACACCGGATTA-----GTACCTGATATTTCTTGACA 2777
QY 549 HisSerSerValAspArgTyrAsnAlaIleSerAspIleIleThrMetIleProAla 568
Db 2778 CATCGAGTGGCGATCTTATTAATGACAGTTCATTCAGATTAATAATTAATCAATTCGCGTC 2837

```

QY 569 ILELYGILYASNAENLEU-----ASPTHRASNSERLYVALLEGU 582  
Db 2838 GTPAAGGTTTGATTTGGCTCCCTATACAGAGGCGCAATATACCTTGATCG 2897  
QY 583 GTPROGILYHISTHCDLYGLYANLEUVALTYRLEUHLNLSERGLNGLY----- 598  
Db 2898 GGTCTCGAATTTACAGAGGAGGAGTAATAAGATATAGAAATGAGTAATATATATCA 2957  
QY 599 -----ARGLEUGLUILETHRCYSGIUTHRPROASNSERTHRCLNSERTHRPHEILEARG 616  
Db 2958 CATATCGGTTAAATTTCA-----GACATTAACAATAATATAGATATAGAG 3005  
QY 617 LEUATGYTATAThRANGLYALAGLYAENTHYLEUPROASNLLESERLEUTHRILEPRO 636  
Db 3006 ATTCGGTATAGCTTCC-----GCTATATATAGTAATTTATATATATCTCTGAAAGA 3059  
QY 637 GLYVALILEGLYTLERPROGILNARGLEUASNAENTHRPHESEGLYTHRANTYRAN 656  
Db 3060 AACGTTAAATCT--CACGCTCAAAAACATGATGATGAGTGAAAGCTTTAACATATAT 3116  
QY 657 AENLEUGLINTYRGLYASPHREGIYTYRPHGLNPHROSERTHR-----VALTHR 673  
Db 3117 AATTTAATATATGCGACTTTGCCCTATTAATTTACGACAAACGACCTTTCATTAAT 3176  
QY 674 LEUPROLEUASNAARGANILEPROPHLEIPHEASNAARGALASPVALSERANSERILE 693  
Db 3177 CTAGGGGCT-----ATATTTGAAGCGAAGACTTCTTGGAATTTGA 3218  
QY 694 LEUILEILASPVYIILEGLUPHEILEPROILETHRSERSEMECHISELINASNAAGLU 713  
Db 3219 GCTTATATATGACCGAATTCGAATTTATCCAGTATGATGAGACATAT-----GAAAGCG 3269  
QY 714 LYSGLNLYLEUGLINTHRILLEGINTHRYLLEAENTHRPHEPHETHRAN 730  
Db 3270 GAACAAGATTTAGAACACGAGCAAGAAAGACGATGATGCTTGTTCAGAT 3320

## RESULT 15

US-10-606-320-18  
Sequence 18, Application US/10606320  
Publication No. US20040091505A1

## GENERAL INFORMATION:

APPLICANT: Andre R. Abad  
APPLICANT: Ronald D. Flannagan  
APPLICANT: Rafael Herrmann  
APPLICANT: Theodore W. Kahn  
APPLICANT: Albert L. Lu  
APPLICANT: Billy Fred McCutchen  
APPLICANT: James K. Presnall  
APPLICANT: James F.H. Wong  
APPLICANT: Cao-Guo Yu  
TITLE OF INVENTION: Genes Encoding Proteins With Pesticidal  
FILE OF INVENTION: Activity  
FILE REFERENCE: 35718/263948  
CURRENT APPLICATION NUMBER: US/10/606,320  
PRIOR FILING DATE: 2003-06-25  
PRIOR APPLICATION NUMBER: 60/391,786  
PRIOR FILING DATE: 2002-06-26  
PRIOR APPLICATION NUMBER: 60/460,787  
PRIOR FILING DATE: 2003-04-04  
NUMBER OF SEQ ID NOS: 134  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 18  
LENGTH: 6613  
TYPE: DNA  
ORGANISM: Bacillus thuringiensis  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (0)...(0)  
OTHER INFORMATION: Genomic Cyt1218-2  
US-10-606-320-18

Alignment Scores:  
Pred. No.:

1.47e-63

Length:

6613

Score: 703.00 Matches: 219  
Percent Similarity: 44.5% Conservative: 127  
Best Local Similarity: 28.2% Mismatches: 289  
Query Match: 18.2% Indels: 142  
DB: 7 Gaps: 32

US-10-783-417-2 (1-735) x US-10-606-320-18 (1-6613)

QY 5 AENASNAENGLUTYRGLUILEILEAPSERHISTHRSEPROTYRPHROASNARG 24  
Db 1263 AATATCAAAATGATATGAAATTTATATAGATGACACCTTCT-----ACTTCT 1310  
QY 25 AENSERASNAENPSE--ARGTYRPROTYRTHRAENAPROASNLNPROLEUGLNAEN 43  
Db 1311 GATACCAATGATTTCTACAGATACCCCTTTGGAAATGAGCCCAAAATGCGTTACAAAAT 1370  
QY 44 THRANTYRLYSGILNTRPLEUSMETCYSGINGLYAEN---THRGINTYRGLYASPAEN 62  
Db 1371 ATGATATATAAAGATTTATTAATAATGCTCGGGAATAATGCTATGAAATACCTGGTTCA 1430  
QY 63 PHEGLUTHRPHEALASERLALASPTHRILALALVALSERALAGLYTHRILEVALSER 82  
Db 1431 CCGTAAGATCTTGTTAGCGGACAAAGAT-----GACGCTAAGCGCGCAATATATATAGTA 1484  
QY 83 GLYTHRLEUVALAGLYTLLEGIYLYLEUTHRSERILESERGLYPROILEGLYILEIE 102  
Db 1485 GGTAAATTTACTATCAAGTTTAGG-----GTCCCATTTGTT 1520  
QY 103 GLYVALILEILEISERPHREGIYTYRLEUILETHRVALPHETRPROALAGLYGLUGIN 122  
Db 1521 GGGCGGATGATGATCTTTATATCTCAACTATTTATGATTTCTGTCGCTTACAGGCA-- 1577  
QY 123 AEPYSTRVALTYRTHCDINPHEILEHLEYSMETGLYLUILEPHEVALASPTHRPROLEU 142  
Db 1578 ---AAGATCATATGGAAGATTTTATGCAACAAGATGAAAGATCTGATTAATCAAAAAATA 1634  
QY 143 THRGUSERILELYSGINLEULYLEUGLINTHRLSEUGLNGLYPHEARGGLINLEUGIN 162  
Db 1635 GCAGATATATGCAAGGAATTAAGCGCTTCCGAAATTAGAAGATATGATTAATATACCA 1694  
QY 163 SERTYRANTHRALALEUASPAEPTTPARGYLSLEULYARGLEUHLNLSERPROGLYLEU 182  
Db 1695 TTTATCTTAACCTGCGCTTGAAAGATGAAAGAA----- 1727  
QY 183 PROPOSERSERLILEUGLNGLNALEALAEUTHRLEULYLEALARGPHEGLUBENVAL 202  
Db 1728 AATCCAAATGTTCA-----AGAGCTTACAGATGTCGAAATTCGATTAATATCCG 1781  
QY 203 HISAENASPHLEILARGGLUILEPROGLYPHEGLNLEUGLINTHRYTRYRTHRLEUEN 222  
Db 1782 GATAGTTATTTTACCAATATACATGCTTTTCCAGTATGACAAATTTGAAATACCATTC 1841  
QY 223 LEUPROILETYRATGLNALALANPHEHLSLEUASLEULEUGLNGLYALAGLU 242  
Db 1842 CTTACAGATATATACAGGACGACCACTTCAATTTACTGTATTTAAAGACCTTCAATT 1901  
QY 243 LEUALAEPGLUTTPASNALASPTLHISPROSERGLINLEGLUPROASNALAGLYTHR 262  
Db 1902 TTTGAGAAAGATGGGATG-----TCTAACACCAT 1934  
QY 263 SERASPAEPTTYRILYLSLEULEULYSGLUANILEPROLYTRYRSEANTYCYVALA 282  
Db 1935 ATTAATATCTATTTATGATGCTCAATATGAAATTTACTGACAGATATTTCTGATCATGTGTA 1994  
QY 283 AENTHRYTRYRATGTHGLYLEULYASLEULYASPAEGLUPROASMETLYSTRPSEITIE 302  
Db 1995 AAGTGATGAAACCTGTTTACCAAAATTTAAAGCACAGAGCGCTTAACATGGGTGAC 2054  
QY 303 PHEASNAEPTYRATGYRMEETHRIETHRVALLLEUASPTHRILESERGLNPHESER 322  
Db 2055 TATTAACCAATTCGATGAGAAATGACATGACGCTTTTATGATGTTGTTGATTAATTCCTCA 2114  
QY 323 LEUTYRASPILLEYAARGTYRARGASPSERILEGLIYLYLEGLUVALYSGLYILEY 342

Db	2115	AATTATGACACACCGCAGTAC-----	CCAATGGAACGAAAGCA-----	2153
Qy	343	AsnGluLeuThrArgGluIleTyrThrGluIleAsnPheAspArgLeuProGluIleu	362	
Db	2154	---CACTAACAAGGAAAGTATATCA-----	GATCCACTGGGGCGGTA 2195	
Qy	363	ArgValGln-----	ProAsnLeuAlaThrMetGluTyr 373	
Db	2196	AACGTCCTTCAATTGTTCTCGTATGACAAGACACCTCTTTCGGAGTAAATGAATCA	2255	
Qy	374	AsnLeuThrArgAlaSerPheLeuPheSerPheLeuGluGlnPheIlePheTyrThr	393	
Db	2256	TCCGTTATTCCAGCACCC---CAGTATTTGATATTAACCGGACCTCACAGTGTATCA	2312	
Qy	394	GluAsnThrAsnPheGlyAsn-----	ArgLeuValGlyIleSerAsnArg 408	
Db	2313	CAATCAAAAGACATTTCTTCGCTCCGATATTAAGACATTTGGCGTGCATCAAAATAGC	2372	
Qy	409	AspAlaProThrTyrSerAsnThrIleThrGluThrLeuTyrGlyValArgThrGlySer	428	
Db	2373	TATCATCGGATTTTATGATTAATTTATTAACAGATGATGAACTTAACAAATCTA	2432	
Qy	429	ProThrThrIlyeThrIleArgProPheGluSerTyr-----	Lys 441	
Db	2433	CACAGCACTAATGACCTTT---GATTTTACGAATTAATGATTTTACACAGCGTTATCAAA	2489	
Qy	442	ValSerIleValIleThrAspArgGlnSerProValSerProIleGlnProHisPheIle	461	
Db	2480	GATGGGGTCTCCTTGATATTTGTTTCTCGTATACGTATATN-----TTTTTGGCA	2543	
Qy	462	IleAsnGlnIleGluLeuTyrLeuAsnGlySerSerAsnAsn-----	ThrLeuLys 478	
Db	2544	ATGCCAGAAAGTCAGGTTTTTCATGGTAAACCAATTGATTAATACAGAAACGTTAAAG	2603	
Qy	479	TyrSer-----	AlaGlyIleSerLeuSerAsnThrGlnAsnThr 491	
Db	2604	TATATCCGGTTCCAAAGATATTATATACGGGG-----	2636	
Qy	492	ThrPhePheGlnPheProArgLysIleAspCysAsnLeuValIleAspProGlyCysSer	511	
Db	2637	-----ACAAAGATTTCCGAATTTAGATTAATCTCCGAAACTTCA	2675	
Qy	512	-----ProAsnPheAsnAsnTyrSerHisIleLeuSerHisPhe---SerLeuPheThr	528	
Db	2676	GATCAACCAATTAATAGATCATATAGCATATGATTAATGATACACAAAGATTTCCGGCG	2735	
Qy	529	TyrSerTyrValIleGlyLeuGlnLeuGlnIleLeuAspThrGlyValLeuGlyTyrThr	548	
Db	2736	ACGGGTTCACTACCGCGATTA-----	GTACCGTATTTTCTTGGACCA 2777	
Qy	549	HisSerSerValAspArgTyrAsnAlaIleSerAspLysIleIleThrMetIleProAla	568	
Db	2778	CATCGAATGCCGAGCTTATTAATGCAATTCATTCAGATAAAATTAATCTCAGATTTCCGCTC	2837	
Qy	569	IleIysGlyAsnAsnLeu-----	AspThrAsnSerLysValIleGlu 582	
Db	2838	GTAAGGTTTCTGATTTGGCTCCCTCTATTAACGAGGGCCAAATTAATACCGTTGTATCG	2897	
Qy	583	GlyProGlyHisIleThrGlyIleAsnLeuValTyrLeuGlnSerGlnGly-----	598	
Db	2898	GGCTCGATTTTACAGGGGGGGGATATAAAGATAAAGAAATGAGATATATATCA	2957	
Qy	599	-----ArgLeuGlnIleThrCysGluThrProAsnSerThrGlnSerTyrPheIleArg	616	
Db	2958	CATATGCGTGTAAAAATTTCA-----	GACATTAACAAAGAAATTAATGATGAGG 3005	
Qy	617	LeuArgTyrAlaIleThrAsnGlyAlaGlyAsnThrLeuProAsnIleSerLeuThrIlePro	636	
Db	3006	ATTGGTATGCTTC-----GCTATATATATCGAATTTATATTAATACCTTTCGAGAA	3059	
Qy	637	GlyValIleGlyIleProGlnArgLeuAsnAsnThrPheSerGlyThrAsnTyrAsn	656	

Db 3060 AACCTTAATCT---CAGCGCTCAAAAACATGATGAATGAGTGAAGCTTAACTATATAT 3116

Qy 657 AsnLeuGlnTyrgIaAephegIyTyrPheGlnPheProSerThr-----ValThr 673

Db 3117 AAAATTATATATATGCAAGCTTTGCCCTTCATTAAATTTTTCGACACGGAACCTTTCATTACT 3176

Qy 674 LeuProuleuAsnArgAsnIleProPheIlePheAsnArgIaAepValSerAsnSerIle 693

Db 3177 CTAGGGGCT-----ATATTTGAACCGAAGACTTTCTTGGAATTGAA 3218

Qy 694 LeuIleIleAspIySIIeGluPheIleProIleThrSerSerMetHisGlnAsnArgGlu 713

Db 3219 GCTTATATATAGCCGAATCGAATTTATATCCCATGATATGACATAT-----GAAGCG 3269

Qy 714 LysGlnIynIySleuGlnIuThrIleGlnThrLysIleAsnThrPhePheThrAsn 730

Db 3270 GAACAGAGATTTAGAAGCAGCGACGAAAGACAGTGAATGCTTTGTTAGCAAT 3320

Search completed: January 20, 2006, 22:58:33  
Job time : 1526 secs

**THIS PAGE BLANK (US#T0)**

GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - nucleic search, using frame\_plus.p2n model

Run on: January 20, 2006, 17:08:39 ; Search time 334 Seconds  
(without alignments)  
3911.700 Million cell updates/sec

Title: US-10-783-417-2

Perfect score: 3869  
Sequence: 1 MNQNNNDNNEYIIDSHTSPY.....KLETIQTQKINTEFTHHTKTL 735

Scoring table:

BLOSUM62	
Xgapop 10.0 , Xgapext 0.5	
Ygapop 10.0 , Ygapext 0.5	
Fgapop 6.0 , Fgapext 7.0	
Delop 6.0 , Delext 7.0	

Searched: 1303057 seqs, 888780828 residues

Total number of hits satisfying chosen parameters: 2606114

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODE=frame+g2n.model -DEV=xlh  
-O=/cg2n\_1/USPTO\_spool/US10783417/runat\_20012006\_095011\_23893/app\_query.fasta\_1.903  
-DB=Issued Patents NA -QFMT=fastp -SUFFIX=g2n.rml -MINMATCH=0.1 -LOOPCL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=biosum62 -TRANS=humana0.cdi  
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=pct -NORM=ext -HBASESIZE=500 -MINLEN=0 -MAXLEN=200000000  
-USER=US10783417@cgn\_1\_1\_200 @runat\_20012006\_095011\_23893 -NCPU=6 -ICPU=3  
-NO MMAP -NRG SCORES=0 -WAIT -DSPBLOCK=100 -DEV TIMEOUT=120 -WARN TIMEOUT=30  
-THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7 -YGAPOP=10  
-YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Issued Patents\_NA:\*  
1: /cg2n\_6/prodata/1/ina/1-COMB.seq:\*  
2: /cg2n\_6/prodata/1/ina/5-COMB.seq:\*  
3: /cg2n\_6/prodata/1/ina/6A-COMB.seq:\*  
4: /cg2n\_6/prodata/1/ina/6B-COMB.seq:\*  
5: /cg2n\_6/prodata/1/ina/H-COMB.seq:\*  
6: /cg2n\_6/prodata/1/ina/PCTUS-COMB.seq:\*  
7: /cg2n\_6/prodata/1/ina/PP-COMB.seq:\*  
8: /cg2n\_6/prodata/1/ina/RE-COMB.seq:\*  
9: /cg2n\_6/prodata/1/ina/Backfile1.seq:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1060.5	27.4	3543	3	US-09-224-024-27
2	1060.5	27.4	3543	3	PCT-US94-07902-27
3	899.5	23.2	2061	3	US-09-224-024-30
4	899.5	23.2	2061	6	PCT-US94-07902-30
5	775.5	20.0	3507	2	US-08-915-468-3
6	676	17.5	1953	2	US-08-915-468-5
7	676	17.5	1953	3	US-07-941-650A-3
8	663.5	17.1	2430	6	PCT-US92-00040-1
9	663	17.1	1959	3	US-08-996-441B-37

10	663	17.1	1959	3	US-08-993-722A-37	Sequence 37, Appl
11	663 <td>17.1</td> <td>1959<td>3<th>US-08-993-170A-37</th><th>Sequence 37, Appl</th></td></td>	17.1	1959 <td>3<th>US-08-993-170A-37</th><th>Sequence 37, Appl</th></td>	3 <th>US-08-993-170A-37</th> <th>Sequence 37, Appl</th>	US-08-993-170A-37	Sequence 37, Appl
12	663 <td>17.1</td> <td>1959<td>3<th>US-08-993-775B-37</th><th>Sequence 37, Appl</th></td></td>	17.1	1959 <td>3<th>US-08-993-775B-37</th><th>Sequence 37, Appl</th></td>	3 <th>US-08-993-775B-37</th> <th>Sequence 37, Appl</th>	US-08-993-775B-37	Sequence 37, Appl
13	663 <td>17.1</td> <td>1959<td>3<th>US-09-427-770-37</th><th>Sequence 37, Appl</th></td></td>	17.1	1959 <td>3<th>US-09-427-770-37</th><th>Sequence 37, Appl</th></td>	3 <th>US-09-427-770-37</th> <th>Sequence 37, Appl</th>	US-09-427-770-37	Sequence 37, Appl
14	663 <td>17.1</td> <td>1959<td>3<th>US-09-427-769-37</th><th>Sequence 37, Appl</th></td></td>	17.1	1959 <td>3<th>US-09-427-769-37</th><th>Sequence 37, Appl</th></td>	3 <th>US-09-427-769-37</th> <th>Sequence 37, Appl</th>	US-09-427-769-37	Sequence 37, Appl
15	660.5 <td>17.1</td> <td>1959<td>3<th>US-07-973-320-3</th><th>Sequence 3, Appl1</th></td></td>	17.1	1959 <td>3<th>US-07-973-320-3</th><th>Sequence 3, Appl1</th></td>	3 <th>US-07-973-320-3</th> <th>Sequence 3, Appl1</th>	US-07-973-320-3	Sequence 3, Appl1
16	660 <td>17.1</td> <td>1959<td>3<th>US-09-377-466B-3</th><th>Sequence 3, Appl1</th></td></td>	17.1	1959 <td>3<th>US-09-377-466B-3</th><th>Sequence 3, Appl1</th></td>	3 <th>US-09-377-466B-3</th> <th>Sequence 3, Appl1</th>	US-09-377-466B-3	Sequence 3, Appl1
17	660 <td>17.1</td> <td>1959<td>3<th>US-10-232-665-3</th><th>Sequence 3, Appl1</th></td></td>	17.1	1959 <td>3<th>US-10-232-665-3</th><th>Sequence 3, Appl1</th></td>	3 <th>US-10-232-665-3</th> <th>Sequence 3, Appl1</th>	US-10-232-665-3	Sequence 3, Appl1
18	658 <td>17.0</td> <td>1959<td>3<th>US-08-996-441B-37</th><th>Sequence 27, Appl</th></td></td>	17.0	1959 <td>3<th>US-08-996-441B-37</th><th>Sequence 27, Appl</th></td>	3 <th>US-08-996-441B-37</th> <th>Sequence 27, Appl</th>	US-08-996-441B-37	Sequence 27, Appl
19	658 <td>17.0</td> <td>1959<td>3<th>US-08-996-441B-37</th><th>Sequence 27, Appl</th></td></td>	17.0	1959 <td>3<th>US-08-996-441B-37</th><th>Sequence 27, Appl</th></td>	3 <th>US-08-996-441B-37</th> <th>Sequence 27, Appl</th>	US-08-996-441B-37	Sequence 27, Appl
20	658 <td>17.0</td> <td>1959<td>3<th>US-08-993-722A-27</th><th>Sequence 27, Appl</th></td></td>	17.0	1959 <td>3<th>US-08-993-722A-27</th><th>Sequence 27, Appl</th></td>	3 <th>US-08-993-722A-27</th> <th>Sequence 27, Appl</th>	US-08-993-722A-27	Sequence 27, Appl
21	658 <td>17.0</td> <td>1959<td>3<th>US-08-993-722A-35</th><th>Sequence 35, Appl</th></td></td>	17.0	1959 <td>3<th>US-08-993-722A-35</th><th>Sequence 35, Appl</th></td>	3 <th>US-08-993-722A-35</th> <th>Sequence 35, Appl</th>	US-08-993-722A-35	Sequence 35, Appl
22	658 <td>17.0</td> <td>1959<td>3<th>US-08-993-170A-27</th><th>Sequence 27, Appl</th></td></td>	17.0	1959 <td>3<th>US-08-993-170A-27</th><th>Sequence 27, Appl</th></td>	3 <th>US-08-993-170A-27</th> <th>Sequence 27, Appl</th>	US-08-993-170A-27	Sequence 27, Appl
23	658 <td>17.0</td> <td>1959<td>3<th>US-08-993-170A-35</th><th>Sequence 35, Appl</th></td></td>	17.0	1959 <td>3<th>US-08-993-170A-35</th><th>Sequence 35, Appl</th></td>	3 <th>US-08-993-170A-35</th> <th>Sequence 35, Appl</th>	US-08-993-170A-35	Sequence 35, Appl
24	658 <td>17.0</td> <td>1959<td>3<th>US-08-993-775B-37</th><th>Sequence 37, Appl</th></td></td>	17.0	1959 <td>3<th>US-08-993-775B-37</th><th>Sequence 37, Appl</th></td>	3 <th>US-08-993-775B-37</th> <th>Sequence 37, Appl</th>	US-08-993-775B-37	Sequence 37, Appl
25	658 <td>17.0</td> <td>1959<td>3<th>US-08-993-775B-35</th><th>Sequence 35, Appl</th></td></td>	17.0	1959 <td>3<th>US-08-993-775B-35</th><th>Sequence 35, Appl</th></td>	3 <th>US-08-993-775B-35</th> <th>Sequence 35, Appl</th>	US-08-993-775B-35	Sequence 35, Appl
26	658 <td>17.0</td> <td>1959<td>3<th>US-09-427-770-27</th><th>Sequence 27, Appl</th></td></td>	17.0	1959 <td>3<th>US-09-427-770-27</th><th>Sequence 27, Appl</th></td>	3 <th>US-09-427-770-27</th> <th>Sequence 27, Appl</th>	US-09-427-770-27	Sequence 27, Appl
27	658 <td>17.0</td> <td>1959<td>3<th>US-09-427-770-35</th><th>Sequence 35, Appl</th></td></td>	17.0	1959 <td>3<th>US-09-427-770-35</th><th>Sequence 35, Appl</th></td>	3 <th>US-09-427-770-35</th> <th>Sequence 35, Appl</th>	US-09-427-770-35	Sequence 35, Appl
28	658 <td>17.0</td> <td>1959<td>3<th>US-09-427-769-27</th><th>Sequence 27, Appl</th></td></td>	17.0	1959 <td>3<th>US-09-427-769-27</th><th>Sequence 27, Appl</th></td>	3 <th>US-09-427-769-27</th> <th>Sequence 27, Appl</th>	US-09-427-769-27	Sequence 27, Appl
29	658 <td>17.0</td> <td>1959<td>3<th>US-09-427-769-35</th><th>Sequence 35, Appl</th></td></td>	17.0	1959 <td>3<th>US-09-427-769-35</th><th>Sequence 35, Appl</th></td>	3 <th>US-09-427-769-35</th> <th>Sequence 35, Appl</th>	US-09-427-769-35	Sequence 35, Appl
30	656 <td>17.0</td> <td>1962<td>3<th>US-09-377-466B-5</th><th>Sequence 5, Appl1</th></td></td>	17.0	1962 <td>3<th>US-09-377-466B-5</th><th>Sequence 5, Appl1</th></td>	3 <th>US-09-377-466B-5</th> <th>Sequence 5, Appl1</th>	US-09-377-466B-5	Sequence 5, Appl1
31	656 <td>17.0</td> <td>1962<td>3<th>US-10-232-665-5</th><th>Sequence 5, Appl1</th></td></td>	17.0	1962 <td>3<th>US-10-232-665-5</th><th>Sequence 5, Appl1</th></td>	3 <th>US-10-232-665-5</th> <th>Sequence 5, Appl1</th>	US-10-232-665-5	Sequence 5, Appl1
32	656 <td>17.0</td> <td>2050<td>3<th>US-08-996-441B-101</th><th>Sequence 101, App</th></td></td>	17.0	2050 <td>3<th>US-08-996-441B-101</th><th>Sequence 101, App</th></td>	3 <th>US-08-996-441B-101</th> <th>Sequence 101, App</th>	US-08-996-441B-101	Sequence 101, App
33	656 <td>17.0</td> <td>2050<td>3<th>US-08-993-722A-101</th><th>Sequence 101, App</th></td></td>	17.0	2050 <td>3<th>US-08-993-722A-101</th><th>Sequence 101, App</th></td>	3 <th>US-08-993-722A-101</th> <th>Sequence 101, App</th>	US-08-993-722A-101	Sequence 101, App
34	656 <td>17.0</td> <td>2050<td>3<th>US-08-993-170A-101</th><th>Sequence 101, App</th></td></td>	17.0	2050 <td>3<th>US-08-993-170A-101</th><th>Sequence 101, App</th></td>	3 <th>US-08-993-170A-101</th> <th>Sequence 101, App</th>	US-08-993-170A-101	Sequence 101, App
35	656 <td>17.0</td> <td>2050<td>3<th>US-08-993-775B-101</th><th>Sequence 101, App</th></td></td>	17.0	2050 <td>3<th>US-08-993-775B-101</th><th>Sequence 101, App</th></td>	3 <th>US-08-993-775B-101</th> <th>Sequence 101, App</th>	US-08-993-775B-101	Sequence 101, App
36	656 <td>17.0</td> <td>2050<td>3<th>US-09-427-770-101</th><th>Sequence 101, App</th></td></td>	17.0	2050 <td>3<th>US-09-427-770-101</th><th>Sequence 101, App</th></td>	3 <th>US-09-427-770-101</th> <th>Sequence 101, App</th>	US-09-427-770-101	Sequence 101, App
37	656 <td>17.0</td> <td>2050<td>3<th>US-09-427-769-101</th><th>Sequence 101, App</th></td></td>	17.0	2050 <td>3<th>US-09-427-769-101</th><th>Sequence 101, App</th></td>	3 <th>US-09-427-769-101</th> <th>Sequence 101, App</th>	US-09-427-769-101	Sequence 101, App
38	655 <td>16.9</td> <td>1959<td>3<th>US-08-996-441B-17</th><th>Sequence 17, Appl</th></td></td>	16.9	1959 <td>3<th>US-08-996-441B-17</th><th>Sequence 17, Appl</th></td>	3 <th>US-08-996-441B-17</th> <th>Sequence 17, Appl</th>	US-08-996-441B-17	Sequence 17, Appl
39	655 <td>16.9</td> <td>1959<td>3<th>US-08-996-441B-59</th><th>Sequence 59, Appl</th></td></td>	16.9	1959 <td>3<th>US-08-996-441B-59</th><th>Sequence 59, Appl</th></td>	3 <th>US-08-996-441B-59</th> <th>Sequence 59, Appl</th>	US-08-996-441B-59	Sequence 59, Appl
40	655 <td>16.9</td> <td>1959<td>3<th>US-08-993-722A-17</th><th>Sequence 17, Appl</th></td></td>	16.9	1959 <td>3<th>US-08-993-722A-17</th><th>Sequence 17, Appl</th></td>	3 <th>US-08-993-722A-17</th> <th>Sequence 17, Appl</th>	US-08-993-722A-17	Sequence 17, Appl
41	655 <td>16.9</td> <td>1959<td>3<th>US-08-993-722A-59</th><th>Sequence 59, Appl</th></td></td>	16.9	1959 <td>3<th>US-08-993-722A-59</th><th>Sequence 59, Appl</th></td>	3 <th>US-08-993-722A-59</th> <th>Sequence 59, Appl</th>	US-08-993-722A-59	Sequence 59, Appl
42	655 <td>16.9</td> <td>1959<td>3<th>US-08-993-170A-17</th><th>Sequence 17, Appl</th></td></td>	16.9	1959 <td>3<th>US-08-993-170A-17</th><th>Sequence 17, Appl</th></td>	3 <th>US-08-993-170A-17</th> <th>Sequence 17, Appl</th>	US-08-993-170A-17	Sequence 17, Appl
43	655 <td>16.9</td> <td>1959<td>3<th>US-08-993-170A-59</th><th>Sequence 59, Appl</th></td></td>	16.9	1959 <td>3<th>US-08-993-170A-59</th><th>Sequence 59, Appl</th></td>	3 <th>US-08-993-170A-59</th> <th>Sequence 59, Appl</th>	US-08-993-170A-59	Sequence 59, Appl
44	655 <td>16.9</td> <td>1959<td>3<th>US-08-993-775B-17</th><th>Sequence 17, Appl</th></td></td>	16.9	1959 <td>3<th>US-08-993-775B-17</th><th>Sequence 17, Appl</th></td>	3 <th>US-08-993-775B-17</th> <th>Sequence 17, Appl</th>	US-08-993-775B-17	Sequence 17, Appl
45	655 <td>16.9</td> <td>1959<td>3<th>US-08-993-775B-59</th><th>Sequence 59, Appl</th></td></td>	16.9	1959 <td>3<th>US-08-993-775B-59</th><th>Sequence 59, Appl</th></td>	3 <th>US-08-993-775B-59</th> <th>Sequence 59, Appl</th>	US-08-993-775B-59	Sequence 59, Appl

#### ALIGNMENTS

RESULT 1  
US-09-224-024-27  
Sequence 27, Application US/09224024  
Patent No. 6056953  
GENERAL INFORMATION:  
APPLICANT: Leslie Hickie  
APPLICANT: Jewel Payne  
TITLE OF INVENTION: Materials and Methods for the Control of  
TITLE OF INVENTION: Calliphoridae Pests  
NUMBER OF SEQUENCES: 31  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: David R. Saliwanhik  
STREET: 2421 N.W. 41st Street, Suite A-1  
CITY: Gainesville  
STATE: FL  
COUNTRY: USA  
ZIP: 32606  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/224,024  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/856,226  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Saliwanhik, David R.  
REGISTRATION NUMBER: 31,794  
REFERENCE/DOCKET NUMBER: MAY9

TELECOMMUNICATION INFORMATION:  
TELEPHONE: 904-375-8100  
TELEFAX: 904-372-5800  
INFORMATION FOR SEQ ID NO: 27:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 3543 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
US-09-224-024-27

Alignment Scores:  
Pred. No.: 2,546-102 Length: 3543  
Score: 1060.50 Matches: 271  
Percent Similarity: 50.3% Conservative: 117  
Best Local Similarity: 35.1% Mismatches: 279  
Query Match: 27.4% Indels: 105  
Gaps: 24

US-10-783-417-2 (1-735) x US-09-224-024-27 (1-3543)

```
QY      1 MetAaGlnAaAsnAspAsnAsnGluTyrGluIleIleAspSerHisThrSerProTyr 20
Db      1 ATGAATCCTTATCAAAATTAATAATGAAATGAAACATTAAATGCTTCAACAAAAAATTA 60
QY      21 PheProAaArgAaSerAsnAspSerArgTyrProTyrThrAsnAsnProAsnGlnPro 40
Db      61 -----AATATATCTTAATATATATACAAAGATATCCAAATGAAATATGTCGCAAAACATTA 114
QY      41 LeuGlnAaThrAsnTyrIleGluTyrLeuAaMetCysGlnGlyAsnThrGlnTyrGly 60
Db      115 TTACAAAGATACAAATTAATAAGATGGCTCATATGTCGCAACAAATCAGCAGATAGT 174
QY      61 AspaAsnGluThrPheAlaSerAlaAspThrIleAlaIleValSerIleGlyThrIle 80
Db      175 GGAGATTTTGAACCTTTATGATGATG-----GGTGAACCTAGTGCCTATATCTAT 225
QY      81 ValSerGlyThrLeuAlaGlyIleGlyLeuThrSerIleSerGlyProIleGly 100
Db      226 GTAGTTGGGACCGTACTGACTGTTCCGGTTCACAAAC-----CCCTTAGA 273
QY      101 IleIleGlyAlaIleIleIleSerPheGlyThrLeuIleThrValPheThrProAlaGly 120
Db      274 CTT-----GCTTAAATAGTTTGTGTCATTAATACAGTCTTTTCCAGCCCA 324
QY      121 GluGlnAspLysThrValTyrThrGlnPheIleLysMetGlyGluIlePheValAspThr 140
Db      325 GACCACATCTAACACA--TGAGTGACTTTATTAACACAAACTAAATAATTTATTAATAAAA 381
QY      141 ProLeuThrGluSerIleGlnLeuLysLeuGlnThrLeuGluGlyPheArgGlnIle 160
Db      382 GAAATAGCATCAACATATATAGTATGCTAATTAATAATTTTAAACAGTCGTTAATGTT 441
QY      161 LeuGlnSerTyrAsnThrAlaLeuAspAspTyrPargLysLeuLysArgLeuGlnAlaPro 180
Db      442 ATCAGCACTTATCATATCACTTAATAAATCAGGAG-----AATATATCA 486
QY      181 GlyLeuProProSerSerAlaLeuGlnGlnAlaIleLeuThrLeuLysIleArgPheGlu 200
Db      487 AACCCACAAATAATCTCAGATGTATAGACACAAATCCAGCTAGTTCATTATCCATTTT 546
QY      201 AsnValHisAsnAspPheIleArgGluIlePro-----GlyPheGlnLeuGlnThrTyr 218
Db      547 AATGTCATTTCCAGAGCTGTAACTCTTGTCTCTTAATCTTAGATGTTCCGATTAATCTAT 606
QY      219 LysThrLeuLeuLeuProIleTyrAlaGlnAlaAlaAsnPheHisLeuAsnLeuGln 238
Db      607 AACATATAGTATATCTAGTATATGACACAGCAAGAACTTATCATCTGATCTATTAAT 666
QY      239 GlnGlyAlaGluLeuAlaAspGluTyrPheAsnAlaAspIleHisProSerGlnIleGluPro 258
Db      667 CAAGCCGTCAAATTTGAAGCGTATTTAAACAAATCAATTCATTCATTTATAGAGCT 726
```

```
QY      259 AsnAlaGlyThrSerAspAspTyrTyrIleLeuLeuGluAsnIleProLysTyrSer 278
Db      727 ---TTGCCACAGCAAGATGATTAATATTCAGATTTAGCTAAAGCATATGAAGATTAACCT 783
QY      279 AsnTyrCysAlaAsnThrTyrArgThrGlyLeuLysAsnLeuArgAspGluPro----- 296
Db      784 AATATATGTGTACCAACTTATTAATAAGATTTAAATTAATAAACAACCGCTGATAGT 843
QY      297 -----AsnMetLysTrpSerIlePheAsnAspTyrArgArgTyrMetThrIle 312
Db      844 AATCTTGATGGAATATATAACTGGAACATACATACCTATTCGAACAAAATGACTACT 903
QY      313 ThrValLeuAspThrIleSerGlnPheSerLeuTyrAspIleLysArgTyrArgAspSer 332
Db      904 GCTGATATAGATCTTGTCGACTCTTCTTAATATGATGATGATTAAT----- 954
QY      333 IleGlyGlyIleGluValLysGlyIleLysAsnGluLeuThrArgGluIleTyrThrThr 352
Db      955 -----CCATAGGTGTCCAAATCTGAATCTGAACTTACCGAATATTTATCAGGTA 999
QY      353 GluIleAsnPheAspArgLeuProGlnLeuArgValGlnProAsnLeuAlaThrMetGlu 372
Db      1000 ---CTTAACCTCGAAGAAAGCCCTATTAATATAT-----GACTTCATATATCAAGAG 1050
QY      373 TyrAsnLeuThrArgAlaSerPheLysLeuPheSerPheLeuGluGlnPheIlePheTyr 392
Db      1051 GATTCACCTACACGTAGACCG---CATTTATTTACTTGCTGATTCCTTGAATTTTAT 1107
QY      393 -----ThrGluAsnThrAsnPheGlyAsnArg----- 401
Db      1108 GAAAAAGCGCAACTACTCCATTAATATTTTTCACACACCTTATATATATGTTTCATTAC 1167
QY      402 ---LeuValGlyIleSerAsnArgAspAlaProThrTyrSerAsnThrIleThrGluThr 420
Db      1168 ACACCTGATATATATATCCAAAATCTAGTGTGTTTGGAATTCACAAATGATATCATATAA 1227
QY      421 LeuTyrGlyLysArgThrGlySerProThrThrIleArgProPheGluSerTyr 440
Db      1228 TTA-----AATGCTCTGTTGGCACAATAATATTTATAT----- 1263
QY      441 LysValSerIleValThrAspArgLysSerProProValSerProIleGlnProHisPhe 460
Db      1264 -----TTT 1266
QY      461 IleIleAsnGlnIleGluLeu-----TyrLeuAsnGlySerSerAsnAsnThr--- 476
Db      1267 TTATTAATATGTCATAGCTTATGATATATATTCATTAATGATTTATATATTTAGTAAA 1326
QY      477 -----LeuLysTyrSerAlaGlyGly 483
Db      1327 ATGATTTTTTATATACATATAGTACTAGTCTTTTGAGAAAGAACTTACAGCAGAGACT 1386
QY      484 SerLeuSerAsnTyrGln---AsnThrThrPhePheGlnPheProArgLysLysAspCys 502
Db      1387 GGGCAATATACCTTATGATGATTAATAATATTTTCCGGTATCCAAATCTTAAACCAAGA 1446
QY      503 AsnLeuValIleAspProGlyCysSerProAsnPheAsnAsnTyrSerHisIleLeuSer 522
Db      1447 GAGAAATCAAGAACATCCCTTACCTTTTTCACATATGATTAATAATGTCATATTTTATCA 1506
QY      523 HisPheSerLeuPheThrTyrSerTyrValIleGlyLeuGlnLeuGlnIleLeuAspThr 542
Db      1507 TTTATTAATAAATCTTATAGTATCCCTGCAACATATATAAATCTCAAGGTAT----- 1554
QY      543 GlyValLeuGlyTyrPThrHisSerSerValAspArgTyrAsnAlaIleSerAspLysIle 562
Db      1555 ---ACGTTTGCTGACACACTAGTGTGATCTTAATAAATATCAATTTATATACATTTA 1611
QY      563 IleThrMetIleProAlaIleLysGlyAsnAsnLeuAspThrAsnSerLysValIleGlu 582
Db      1612 ACTACCCAAATTCAGCTGTAAAGCGAATTCATCTGGACCTCTTCTTAAGGTGTTCAA 1671
```



Qy	583	GLYProGlyIh <sup>1</sup> sthrIGLYVAsnleuValTyrleuGlnserInGlyVArGleuGluLe	602
Db	1672	GGACCTGGCTACACGAGGGGGATTAAATT-----GATTCCAAAGATCATTTCAAAATT	1725
Qy	603	ThrCYsgIuThrProAsnSerThrgInserTyrPheIleArgLeuArgTyrAlaThrAsn	622
Db	1726	ACATGTCAACACTCAAAATTTTCAACAAATCGATATTTTATAGAAATTCGTATGCTTCAAAAT	1785
Qy	623	GLYAlaGLYAsnThrleuProAsnIleSerleuThrIleProGlyValIleGlyIlePro	642
Db	1786	GGAAAGCGCAAAATACACGAGCTGTATTAATCTTATATATCCAGGGGTGCAGAATG---	1842
Qy	643	ProGlnArgLeuAsnAsnThrPheSerGlyThraSerTyrAsnAsnLeuGlnTyrGlyAsp	662
Db	1843	GGTATGGCAGCTCAACCCCACTTTTCTGGTACAGATTATAGAAATTTAAATATTAAGAT	1902
Qy	663	PheGlyTyrPheGlnPheProSerThrValThrleuProLeuAsnArgAsnIleProPhe	682
Db	1903	TTTCAGTACTTAGAATTTTCTAACACGAGTGAAATTTGCTCCAAATCAAAACATATCTCTT	1962
Qy	683	IlePheAsnArgAlaAspVal---SerAsnSerIleleuIleIleAspIleIleGluPhe	701
Db	1963	GTGTTTAATCGTTCGAGATGTATATACAAACACACAGATCTATTGATTAATTTGAATTT	2022
Qy	702	IleProIleThrSerSerMetCHisGlnAsnArgGlyIuTyrGlnIuTyrIleGln	721
Db	2023	CTGCCAAATTAATCTCGTTCTTATAGAGAGATAGAGAAACAAAAATTGAAACAGTACA	2082
Qy	722	ThrIuSIIeAsnThrPhePheThrAsnHisThrIys	733
Db	2083	CAAAATTAATTAATCAATTTTATGCAAAATCTATATAAA	2118

```

1 RESULT 2
2 PCT-US94-07902-27
3 Sequence 27, Application PC/TUS9407902
4 GENERAL INFORMATION:
5 APPLICANT:
6 APPLICANT: Street address: 4980 Carroll Canyon Road
7 APPLICANT: City: San Diego
8 APPLICANT: State/Province: California
9 APPLICANT: Country: US
10 APPLICANT: Postal code/Zip: 92121
11 APPLICANT: Phone number: (619) 453-8030 Fax number: (619) 453-6991
12 APPLICANT: Telex number:
13 TITLE OF INVENTION: Materials and Methods for the Control of
14 TITLE OF INVENTION: Calliphoridae Pests
15 NUMBER OF SEQUENCES: 31
16 CORRESPONDENCE ADDRESS:
17 ADDRESSEE: David R. Saliwanchik
18 STREET: 2421 N.W. 41st Street, Suite A-1
19 CITY: Gainesville
20 STATE: FL
21 COUNTRY: USA
22 ZIP: 32606
23 COMPUTER READABLE FORM:
24 MEDIUM TYPE: Floppy disk
25 COMPUTER: IBM PC compatible
26 OPERATING SYSTEM: PC-DOS/MS-DOS
27 SOFTWARE: PatentIn Release #1.0, Version #1.25
28 CURRENT APPLICATION DATA:
29 APPLICATION NUMBER: PCT/US94/07902
30 FILING DATE:
31 CLASSIFICATION:
32 ATTORNEY/AGENT INFORMATION:
33 NAME: Saliwanchik, David R.
34 REGISTRATION NUMBER: 31,794
35 REFERENCE/DOCKET NUMBER: MA79
36 TELECOMMUNICATION INFORMATION:
37 TELEPHONE: 904-375-8100
38 TELEFAX: 904-372-5800
39 INFORMATION FOR SEQ ID NO: 27:
40 SEQUENCE CHARACTERISTICS:
41 LENGTH: 3543 base pairs

```

```

; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
PCT-US94-07902-27

```

### Alignment Scores:

Pred. No.:	2,54e-102	length:	354
Score:	1060.50	Matches:	271
Percent Similarity:	50.3%	Conservative:	117
Best Local Similarity:	35.1%	Mismatches:	279
Query Match:	27.4%	Indels:	105
DB:	6	Gaps:	24

US-10-783-417-2 (1-735) X PCT-US94-07902-27 (1-3543)

Qy	MetAsnGlnAsnAspAspAsnGlnGlyArgGluLeileAspSerHisSerProTyr	20
Dp	1 ATGATACCTTATCAAAATATAAATATGAAACATTAAATGCTTCACAAAAAATTA	60
Qy	21 PheProAsnArgAsnSerAsnAspSerArgTyrProTyrThaAsnAspProAsnGlnPro	40
Dp	61 -----AATATATCTAATATATATACAGATATCCAAATAGAAAATATGCGAAAACATTA	114
Qy	41 LeuGlnAsnThrAsnTyrLysGluTrpLeuAsnMetCysGlnGlyAsnThrGlnTyrGly	60
Dp	115 TTTCAAAGTACAAATTTATTAACATTTGGCTCAATATGTCTCAACGAATACGACGATTAAGT	174
Qy	61 AspAsnPheGluThrPheAlaSerAlaAspThrIleAlaIleValSerAlaGlyTyrIle	80
Dp	175 GGAGATTTGAAACTTTATTTATGATAGT-----GGTAGACTCAAGTCCCTATACATTT	225
Qy	81 ValSerGlyTyrThrLeuLeuAlaGlyTyrIleGlyLeuThrSerIleSerGlyProIleGly	100
Dp	226 GTATGTTGGGACCGTATGACTGATGTTTCGGGTTCACAA-----CCCTTAGA	273
Qy	101 IleIleGlyAlaIleIleIleSerPheGlyThrLeuIleThrValPheTrpProAlaGly	120
Dp	274 CTT-----GCTTAAATAGGTTTGTGATTAATTAACAGATTTCTTTTCCAGGCCAA	324
Qy	121 GluGlnAspLysThrValTyrTrpGlnPheIleLysMetGlyGluIlePheValAspThr	140
Dp	325 GACCATCTTAACAC-----TGAGTACCTTTATTAACACAACTAATAATATTATAAAAAA	381
Qy	141 ProLeuThrGluSerIleLysGlnLeuLysLeuGlnThrLeuGlnGlyPheAspGlnIle	160
Dp	382 GAATATGCATCAACATATATATAGTATCTAATAAATTTTMAACAGTCGTTAATGTT	441
Qy	161 LeuGlnSerTyrAsnThrAlaLeuAspAspTrpArgLysLeuLysValSerLeuGlnAlaPro	180
Dp	442 ATCAGACATTATCATATATCATCTTAAAACATGGGAG-----AATATCCA	488
Qy	181 GlyLeuProPheSerSerAlaLeuGlnGlnAlaAlaLeuThrIleLysIleArgProGlu	200
Dp	487 AACCCCAAAATATCTCAGATGTAAGACAAATATCCAGTATGTCATTACCAATTTCAA	546
Qy	201 AsnValHisAsnAspPheIleArgGluIlePro-----GlyPheGlnLeuGlnTyr	218
Dp	547 AATGTCATTCACAGCGCTGTAAACTCTTCTCTCTTAATCTAGATGATGGATGACTAT	606
Qy	219 LysThrLeuLeuLeuProIleTyrAlaGlnAlaIleAsnPheHisLeuAsnLeuLeuGln	238
Dp	607 AACATCTAGTATATCTAGTTTATTCACACACGACCAAACTTACATCTGACGTATTAAT	666
Qy	239 GlnGlyAlaGluLeuAlaAspGluTrpAsnAlaAspIleHisProSerGlnIleGluPro	258
Dp	667 CAAGCCGTCAAAATTTGAACCGTATTTAAAAAACAATCACAATTCGATTTATTTAGAGCCT	726
Qy	259 AsnAlaGlyThrSerAspAspTyrTyrTyrLysLeuLeuGlnIleProLysTyrSer	278
Dp	727 ---TTGCCAACGACAAATGATTTATTTATTCAGATTTGACTTAAGCTTATGAAAGATTACCT	783
Qy	279 AsnTyrCysAlaAsnThrTyrArgThrGlyLeuLysAsnLeuArgAspGluPro-----	286

```

Db      784 AATTATGTGTGAACAACCTTATATAAAAGATTAAATTTAATTAATAAACAAGCGCTGATAGT 843
Oy      297 -----AsmMetIytrpSerIlePheAenAapTyArgArgTyMetThrIle 312
Db      844 AATCTTGATGAAATATATACTGGAACACATACATAGTATCGAACAAAATAGCTACT 903
Oy      313 ThrValLeuAepThrIleSerGlnPheSerLeuTyArgPileuArgTyArgAapSer 332
Db      904 GCTGATATAGATCTTGCACTCTTCTCATATATGATGATGATGATGATGATGATGAT 954
Oy      333 IleGlyGlyIleGluValIleGlyIleuArgGluLeuThrArgGluIleTyThrThr 352
Db      955 -----CCAATAGTGTCCATCTGAATCTGAACTTACCGAATAATTATCAGGTA 999
Oy      353 GluIleAenPheAapArgLeuProGlnLeuArgValGlnProAenLeuAenIleThrMetGlu 372
Db      1000 -----CTTAACCTTGAAAGAAAGCCCTTAATAATTTAT-----GACTTCATATTCAGAG 1050
Oy      373 TyrAenLeuThrArgAlaSerPheLeuPheSerPheLeuGluGlnPheIlePheTy 392
Db      1051 GATTCACCTTACACGTAGACCG---CATTTATTTACTGCTGATTCCTTGATTTTAT 1107
Oy      393 -----ThrGluAenThrAenPheGlyAenArg----- 401
Db      1108 GAAAAAGCGCAACTACTCTCATATTTTTCACACGCCATTATATATGTTTCATTAC 1167
Oy      402 ---LeuValIleGlyIleSerAenArgAapIleProThrTySerAenThrIleThrGluThr 420
Db      1168 ACACCTTGATATATATATCCAAAATCTAGTGTGTTTGGAAATCACAAATGATCTATATA 1227
Oy      421 LeuTyArgIleuArgTyArgTySerProThrThrIleArgProPheGluSerTy 440
Db      1228 TTA-----AATCTCTGTTGGTGGCAACAATATTATTTAT----- 1263
Oy      441 LysValSerIleValThrAapArgIleSerProProValSerProIleGlnProHisPhe 460
Db      1264 ----- 1266
Oy      461 IleIleAenGlnIleGluLeu-----TyrLeuAenGlySerSerAenAenThr--- 476
Db      1267 TTATTAATATGTCATTAAGCTTATAGATTAATAATATCTTAATATGATTAATATTA 1326
Oy      477 -----LeuTySerIleGlyIle 483
Db      1327 ATGATATTTTATATACTAATAGTACTAGACTTTTGGAAAGAAGACTTACACAGAGACT 1386
Oy      484 SerLeuSerAenTyGln---AenThrThrPheIleGlnPheProArgIleLysAapCys 502
Db      1387 GGGCAAAATTAACCTTATGATGATTAATAATATTTTCGGGTTTCCAAATCTTAAACCA 1446
Oy      503 AenLeuValIleAapProGlyCysSerProAenPheAenAenTySerHisIleLeuPhe 522
Db      1447 GAGAAATCAAGCAATGCCCTTACCTTTTCCACATATGATTAACCTATATATTTATCA 1506
Oy      523 HisPheSerLeuPheThrTySerTyValIleGlyLeuGlnLeuGlnIleLeuAapThr 542
Db      1507 TTTATTAATAAGCTTAGTATCCCTGCAACATATAAAGCAAGTGAT----- 1554
Oy      543 GlyValIleuGlyTyThrHisSerSerValAapArgTyArgAalIleSerAapLysIle 562
Db      1555 ---ACGTTTGCTTGACACACTCTAGTGTGATCTCTTAATAAATACATATTATACACAT 1611
Oy      563 IleThrMetIleProAlaIleLeuGlyAenAenLeuAapThrAenSerLysValIleGlu 582
Db      1612 ACTACCCAAATTCAGCTGATTAAGCGAAATTCAGCTGAGCTCTTCAAGGTTCCTCA 1671
Oy      583 GlyProGlyHisThrGlyIleAenLeuValTyLeuGlnSerGlnIleArgLeuGluIle 602
Db      1672 GGAAGCTGGTGCATACGAGAGGAGTTTATTT-----GATTCAAGATCATTTCAAAAT 1725
Oy      603 ThrCysGluThrProAenSerThrGlnSerTyPheIleArgLeuArgTyValAthrAen 622

```

```

Db      1726 ACATGTCACACACTCAAAATTTTCAACATCGTATTTTATTAAGAAATTCGTTAGCTTCAAA 1785
Oy      623 GlyAlaGlyAenThrLeuProAenIleSerLeuThrIleProGlyValIleGlyIlePro 642
Db      1786 GGAAGCGCAAAATACACAGAGCTGTATTAATCTTATGATCCAGGGGTGACAGAACTG--- 1842
Oy      643 ProGlnArgLeuAenAenThrPheSerGlyThrAenTyArgAenAenLeuGlnTyGlyAap 662
Db      1843 GGTATGCGACTCAACCCCACTTTTCTGTCAGATTTATAGAAATTTAAATATATGAAGAT 1902
Oy      663 PheGlyTyThrGlnPheProSerThrValThrLeuProLeuAenAenArgAenIleProPhe 682
Db      1903 TTTCAAGTACTTAAGATTTTCTTAACGAGGTGAATTTGCTCCAAATCAAAACATATCTCTT 1962
Oy      683 IlePheAenArgAlaAapVal---SerAenSerIleLeuIleIleAapLysIleGluPhe 701
Db      1963 GTGTTTATCGTTTCGATGATATATACAAACACACAGACTTATGATTAATAATTTGAATTT 2022
Oy      702 IleProIleThrSerSerMetHisGlnAenArgIleuGlyIleGlyIleGluThrIleGln 721
Db      2023 CTGCCAATTAACCTCGTTCTATTAAGAGAGATGAGAGAAACAAAATTTAGAAAACAGTACAA 2082
Oy      722 ThrLysIleAenThrPhePheThrAenHisThrLys 733
Db      2083 CAATTAATTAATACATTTTATGCAAAATCTATATAAA 2118

RESULT 3
US-09-224-024-30
Sequence 30, Application US/09224024
Patent No. 6056953
GENERAL INFORMATION:
APPLICANT: Leslie Hickie
APPLICANT: Jewel Payne
TITLE OF INVENTION: Materials and Methods for the Control of
TITLE OF INVENTION: Calliphoridae Pests
NUMBER OF SEQUENCES: 31
CORRESPONDENCE ADDRESS:
ADDRESSEE: David R. Saliwanchik
STREET: 2421 N.W. 41st Street, Suite A-1
CITY: Gainesville
STATE: FL
COUNTRY: USA
ZIP: 32606
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/224, 024
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/856,226
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Saliwanchik, David R.
REGISTRATION NUMBER: 31,794
TELEPHONE: 904-375-8100
TELEFAX: 904-375-8100
REFERENCE/DOCKET NUMBER: MA79
TELECOMMUNICATION INFORMATION:
INFORMATION FOR SEQ ID NO: 30:
SEQUENCE CHARACTERISTICS:
LENGTH: 2061 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-09-224-024-30

Alignment Scores: 1.61e-85 Length: 2061
Pred. No.:

```

Score: 899.50 Matches: 247  
Percent Similarity: 48.3% Conservative: 126  
Best Local Similarity: 32.0% Mismatches: 270  
Query Match: 23.2% Indels: 129  
DB: 3 Gaps: 30

US-10-783-417-2 (1-735) x US-09-224-024-30 (1-2061)

1 MetAsnGlnAsnAsnAsnAsnAsnGluTyrGlnIleIleAspSerHisThrSerProTyr 20  
1 ATGATCCATATCAAAATAGATGAAATGAAATATTCATAGCTCATCCAAATGGTTT 60  
21 PheProAsnArgAsnSerAsnAsnSerAspArgTyrProTyrThrAsnAsnProAsnGlnPro 40  
61 -----AGCAAGCTAATACTATTCTAGATATCCATTAGCAAAATAGCAAAATCAACA 114  
41 LeuGlnAsnThrAsnTyrLysGluTyrLeuAsnMetCysGlnGlyAsnThrGlnTyrGly 60  
115 CTGAAATAACAGATATACAAAGATTGGCTCAATGTGTCAGATATATCAACATATATGCG 174  
61 AspAsnPheGluThrPheAlaSerAlaAspThrIleAlaValSerAlaGlyThrIle 80  
175 AATAATGCGGGGAATTTTGTAGTCTGAACATATTGTGAGTATGTCAGGTATATT 234  
81 ValSerGlyThrLeuLeuAlaGlyIleGlyGlyLeuThrSerIleSerGlyProIleGly 100  
235 GTAGTAGGAAGTATGTTAGGA-----GCTTTCCTGCCCCT----- 270  
101 IleIleGlyAlaIleIleIleSerPheGlyThrLeuIleThrValPheTyrProAlaGly 120  
271 GCTTAGTGTGACAGATATATCTTTGGGACTTTTGGCCATCTTTGG--CAAGCA 327  
121 GlnGlnAspLysThrValTyrThrGlnPheIleLysMetGlyGlnIlePheValAspThr 140  
328 TCTGACCTCGCAAAATGTTGGCAGATTGTTAAACATCGCA-----GGAAGC 375  
141 ProLeuThrGlu-----SerIleLysGlnLeuLysLeuGlnThrLeuGlnGlyPhe 157  
376 CCTATTCAGAAATATAGATATAAAACATATTAATATGTAATCTTATCTGTAACCTATA 435  
158 ArgGlnIleLeuGlnSerTyrAsnThrAlaLeuAspAspTyrArgLysLeuLysArgLeu 177  
436 AAAAATCAACCTGATTAATATATCAAGAAATTTTTCGATTAATGGAGCAGACAGCT----- 489  
178 GlnAlaProGlyLeuProProSerSerAlaLeuGlnGlnAlaAlaLeuThrLeuLysIle 197  
490 -----ACACACGCTAATGCTTAAGCAGTACATGATCTCTTACT 528  
198 ArgPheGluAsnValHisAsn-----AspPheIleArgGlnIleProGlyPheGln 214  
529 ACCTTAGAACCTTAATATAGATTAAGATTAGATATGTTAAATAATATGCTAGCTATGCA 588  
215 LeuGluThrTyrIleThrLeuLeuLeuProIleTyrAlaGlnAlaAlaAsnPheHisLeu 234  
589 ATACCAACA-----CTCCGTCATATGCAACAAATAGCTACTTGGCACTTG 633  
235 AsnLeuLeuGlnGlnGlyAlaGluLeuAlaAspGluTyrAsnAlaAspIleHisProSer 254  
634 AATTATTTAAACATGCTGCTACTACTTACATATATATG-----CTGCAAAATCA 684  
255 GlnIleGluProAsnAlaGlyThrSerAspAspTyrTyrLys-----LeuLeuLysGlnLys 273  
685 GGTAATAATCAAGTACTTCAATCATCTAATTAATCTATCAAGGGCTAATTTAAACGTAA 744  
274 IleProLysTyrSerAsnTyrCysAlaAsnThrTyrArgThrGlyLeuLysAsnLeuArg 293  
745 ATACAGATATATCTGACTATTGTATACAAACCTACAGACAGACTAATGATGATTAG 804  
294 AspGluProAsnMetLysTyrSerIlePheAsnAspTyrArgArgTyrMetThrIleThr 313  
805 ACTAATATACCAACCAATGATATATATATATATCTTACCGTTTACAAATGACTCTAATCT 864  
314 ValLeuAspThrIleSerGlnPheSerLeuTyrAspIleLysArgTyrArgAspSerIle 333

865 GTGTAGATCTTATTCCTATTTTCCAAATATATGACCCAGAAATAAT----- 912  
334 GlyGlyIleGluValLysGlyIleLysAsnGlnLeuThrArgGlnIleTyrThrGln 353  
913 -----CCAAATGAGGTAAATCTGAATCTTACCAAGAAAGTTAT--ACGAAT 957  
354 IleAsnPheAspArgLeuProGlnLeuArgValGlnProAsnLeuAlaThrMetGluTyr 373  
958 GTTAATTCAGTACATTT-----AGAACATAAACAGAACTAGAAAT 999  
374 AsnLeuThrArgAlaSerPheLysLeuPheSerPheLeuGlnGlnPheIlePheTyrThr 393  
1000 GGATTAATCTAGA--AATCTCATATTATTACTTGATTAACCAAGGGCGTTTATACACA 1056  
394 GluAsnThrAsn-----PheGlyAsnArgLeu 402  
1057 AGAAATTTTCGAGACATTTGATCTTATGATATTTTCTTTTACAGGTAAACGAGATG 1116  
403 ValGlyIleSerAsnArgAspAlaProThrTyrSerAsnThrIleThrGluThrLeuTyr 422  
1117 -----GCTTTACATCTATATGATGATCGCAACATATCTGG 1155  
423 GlyIuArgThrGly-----SerProThrThrIleArgProPheGluSer 439  
1156 GGAAGCGGTTCATGACATATATTCTTCAAGACATCCAAAGTATTTCTTTTATAGA 1215  
440 Tyr-----LysValSerIleValThrAspArgGlnSerProValSerPro 455  
1216 AACAAACCTATTGATTAAGGTCCAAATTTGCACACATAGACATCTACATATTAATA-- 1272  
456 IleGlnProHisPheIleIleAsnGlnIleGluLeuTyrLeuAsnGlySerSerAsnAsn 475  
1273 -----TATGAAAGATATTTTTCATATGACAGTGA 1305  
476 ThrLeuLysTyrSerAlaGlyLysLeu-----SerAsnTyrGlnAsnThrThrPhePhe 494  
1306 GTATTTCGATATTCATCCAAATTCACAAATAGAAATATATTATTAAGAAGCTGTTAT 1365  
495 GlnPheProArgLysLysLeuArgCysAsnLeuValIleAspProGlyCysSerProAsnPhe 514  
1366 ATGATTTCCAAACAAACATGAAATAAT----- 1392  
515 AsnAsnTyrSerHisIleLeuSerHisPheSerLeuPheThrTyrSerTyrValIleGly 534  
1393 AAAGATATGCTATCTATCTGATATATAAACTGATATATATATTTCAAGATTT 1452  
535 LeuGlnLeuGlnIleLeuAspThrGlyValLeuGlyTyrThrHisSerSerValAspArg 554  
1453 AGAGAAAGAAAGAGCTT-----GCATTATGTCAGACATATCTGTTGATTTTC 1503  
555 TyrAsnAlaIleSerAspLysIleIleThrMetIleProAlaIleLysGlyAsnAsnLeu 574  
1504 CAATATACAAATGATTAATGATTAACATCACCAATATCCAGCTCTAAACCTTTGAAGTA 1563  
575 AspThrAsnSerLysValIleGluGlyProGlnHisThrGlyLysAsnLeuValTyrLeu 594  
1564 AGTTCTGATTTCAAAATTTGTGAAGGTCTGTGCACACGGTGAAGACTGTGAATTTCT 1623  
595 GlnSerGlnGlyArgLeuGluIleThrCysGluThrProAsnSerThrGlnSerTyrPhe 614  
1624 AAAGATAGTATGATTTTATGATTAATTT--TTAAATAATCTTCTGACAAATTCGA 1680  
615 IleArgLeuArgTyrAlaThrAsnGlyAlaGlyAsnThrLeuProAsnIleSerLeuThr 634  
1681 GTACGATATCGTTATGCTACTATGCT-----CCAAAGCAACACAGATATTC 1725  
635 IleProGlyValIleGlyIleProProGlnArgLeuAsnAsnThrPheSerGlyThrAsn 654  
1726 TTACCGGAATGATATCTATTAAGTGTGAG--CTCCAGATGACCACTTCCGCCCAAAAC 1782  
655 TyrAsn-----AsnLeuGlnTyrGlyAspPheGlyTyrPheGlnPheProSerThrVal 672

```

Db      1783 CCAATGCTACGATTTAACATATGCAATTTGGATATGTACATTTCGAAGAACGTT 1842
Qy      673 -----ThrlleuProleuAsnArgAsnIleProphle 683
Db      1843 CCAATTAACATTTGAAGAGAGAGACACTTATTAATGACC-----TTA 1887
Qy      684 PheAsnArgAlaAspValSerAsnSerIleIleIleAspIleIleGluPheIlePro 703
Db      1888 TATGTATACCAATCATTCATATATATA---TATATGACCAAAATCGAATTATTCGA 1944
Qy      704 ILeuThrSerMetHisGlnAsnArgGluLeuGlnLeuGluThrIleGlnThrLys 723
Db      1945 ATCACTCAATCTGTATATAGATTATACAGAGAACCAAAATATAGAAAACACAGAAATA 2004
Qy      724 ILeuThrPhePheThrAsnHisThrLysThrLeu 735
Db      2005 GTGAATGATTATTTGTTAATTAAACAAAGTTCTT 2040

RESULT 4
PCT-US94-07902-30
/ Sequence 30, Application PC/TUS9407902
/ GENERAL INFORMATION:
/ APPLICANT:
/ APPLICANT: Street address: 4980 Carroll Canyon Road
/ APPLICANT: City: San Diego
/ APPLICANT: State/Province: California
/ APPLICANT: Country: US
/ APPLICANT: Postal code/Zip: 92121
/ APPLICANT: Phone number: (619) 453-8030 Fax number: (619) 453-6991
/ APPLICANT: Telex number:
/ TITLE OF INVENTION: Materials and Methods for the Control of
/ TITLE OF INVENTION: Calliphoridae Pests
/ NUMBER OF SEQUENCES: 31
/ CORRESPONDENCE ADDRESSES:
/ ADDRESSES: David R. Saliwanchik
/ STREET: 2421 N.W. 41st Street, Suite A-1
/ CITY: Gainesville
/ STATE: FL
/ COUNTRY: USA
/ ZIP: 32606
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: Patent In Release #1.0, Version #1.25
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: PCT/US94/07902
/ FILING DATE:
/ CLASSIFICATION:
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Saliwanchik, David R.
/ REGISTRATION NUMBER: 31,794
/ REFERENCE/DOCKET NUMBER: MA79
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 904-375-8100
/ TELEFAX: 904-372-5800
/ INFORMATION FOR SEQ ID NO: 30:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 2061 base pairs
/ TYPE: nucleic acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ MOLECULE TYPE: DNA (genomic)
/ PCT-US94-07902-30

Alignment Scores:
Pred. No.: 1.61e-85 Length: 2061
Score: 899.50 Matches: 247
Percent Similarity: 48.3% Conservative: 126
Best Local Similarity: 32.0% Mismatches: 270
Query Match: 23.2% Indels: 129
Db: 6 Gaps: 30

```

```

US-10-783-417-2 (1-735) x PCT-US94-07902-30 (1-2061)
Qy      1 MetAsnGlnAsnAsnAspAsnAsnGluTyrGluIleIleAspSerHisThrSerProTyr 20
Db      1 ATGAATCATATCAAAATTAAGATGAATGAAATATCAATATTCATATGCTCCATCAAGGTTT 60
Qy      21 PheProAsnAlaAsnSerAsnAspSerArgTyrProTyrThrAsnAsnProAsnGlnPro 40
Db      61 -----ACCAAGTCATTAATCAATCTGATTCATATCAATTCAGCAAAATTAACCAATACCA 114
Qy      41 LeuGlnAsnThrAsnTyrLysGluTyrLeuAsnMetCysGlnGlyAsnThrGlnTyrGly 60
Db      115 CTGAAACACACAGCAATTCACAAAGATTGGCTCAATGCTGTCACAGATATCAACAAATATGGC 174
Qy      61 AspAsnPheGluThrPheAlaSerAlaAspThrIleAlaIleValSerAlaGlyThrIle 80
Db      175 AATAATGGGGGAATTTGTTAGTTCTGAACATATGTTGAGTTAGTGAAGGATTAATT 234
Qy      81 ValSerGlyThrLeuLeuAlaGlyIleGlyGlyLeuThrSerIleSerGlyProIleGly 100
Db      235 GTAGTAGGAACTATGTTAGCA-----GCTTTGCTGCCCT----- 270
Qy      101 IleIleGlyAlaIleIleIleSerPheGlyThrLeuIleThrValPheThrProAlaGly 120
Db      271 GCTTAGCTGACGATATATATCTTTGGACCTTGTGGCCGATCTTTGG---CAAGCA 327
Qy      121 GluGlnAspLysThrValTyrThrGlnPheIleLysMetGlyGluIlePheValAspThr 140
Db      328 TGTGACCTGCAAAATGTTTGGACAGATTGTTAAACATCGGA-----GGAAG 375
Qy      141 ProLeuThrGlu-----SerIleLysGlnLeuLysLeuGlnThrLeuGluGlyPhe 157
Db      376 CCTATACAGAAATAGATTAACAAACATTAATGACTTCACTTCACTTAACACCTATA 435
Qy      158 ArgGlnIleLeuGlnSerTyrAsnThrAlaLeuAspAspTyrPargLysLeuLysArgLeu 177
Db      436 AAAAATCAACTGATTAATATCAAGAAATTTTCGATAAATGGAGCCAGCAGC----- 489
Qy      178 GlnAlaProGlyLeuProProSerSerAlaLeuGlnGlnAlaAlaLeuThrLeuLysIle 197
Db      490 -----ACACAGCTAATGCTAATAGCAGTCAATGATCTCTTACT 528
Qy      198 ArgPheGluAsnValHisAsn-----AspPheIleArgGluIleProGlyPheGln 214
Db      529 ACCTTAGAAGCTATATATAGATTAAGATTAGATATGTAATAAATAATGCTAGCATGCA 588
Qy      215 LeuGluThrTyrLysThrLeuLeuProIleTyrAlaGlnAlaAlaAsnPheHisLeu 234
Db      589 ATACCAACA-----CTCCCTGATATGCAAAATAGCTACTTGGACCTTG 633
Qy      235 AsnLeuLeuGlnGlnGlyAlaGluLeuAlaAspGluTyrAsnAlaAspIleHisProSer 254
Db      634 AATTTATTAACATGCGCTACCTATTCATATATATG-----CTGCAAAATCA 684
Qy      255 GlnIleGluProAsnAlaGlyThrSerAspAspTyrTyrLys---LeuLeuLysGluAsn 273
Db      685 GGTATTAATCCAGTACTTCATCAATTCATATATCAATCAAGGGCTATTTAAACGTAA 744
Qy      274 IleProLysTyrSerAsnTyrCysAlaAsnThrTyrArgThrGlyLeuLysAsnLeuArg 293
Db      745 ATACAGAAATATATGACTATTTGTATACAAACGTACAAAGCAGACCTACTATGTATGA 804
Qy      294 AspGluProAsnMetLysTyrSerIlePheAsnAspTyrArgArgTyrMetThrIleThr 313
Db      805 ACTAATACTAACGCAACATGCAATATGTATATCTTACCGTTTGAATATGACTTAAT 864
Qy      314 ValLeuAspThrIleSerGlnPheSerLeuTyrAspIleLysArgTyrArgAspSerIle 333
Db      865 GTGTATAGATCTTATTTGCTATTTTCCAAATATATGACCCAGAAAATAT----- 912
Qy      334 GlyGlyIleGluValLysGlyIleLysAsnGluLeuThrArgGluIleTyrThrThrGlu 353
Db      913 -----CCAATAGAGTTAATCTGAATCTTACAGAGAGTTAT---ACGAAT 957

```

Qy	354	..leAsnPhLeAspAArgLeuProGlnLeuAryValGlnProAsnLeuAlaThrMetGlyIyr	373
Db	958	GTATATTCAGATACATTT-----AGAACCATACACGAACCTGAANAAT	999
Qy	374	AsnLeuThrAArgAlaSerPheIleuPheSerPheLeuGlnGlnPheIlePheThyr	393
Db	1000	GGATTATCTACA---AATCTCACTATTATTTACTGGATAAACCAAGGCGCTTTTACCA	1056
Qy	394	GluAsnThrAsn-----PheGlyAsnArgLeu	402
Db	1057	AGAAATTCGAGACATCTTGATCCTATGATATTTTCTTTTACAGGTAAACAAGATG	1116
Qy	403	ValGlyIleSerAsnArgAspAlaProThrTySerAsnThrIleThrGlnThLeuTy	422
Db	1117	-----GCCCTTTCACACTACTAATGATGATGCGCACTAATATCTGG	1155
Qy	423	GlyGluAArgThrGly-----SerProThrThyStrIleArgProPheGlnSer	439
Db	1156	GGAGCGGCTTCAGGACATATTATTTCTCAAGACATACCAAGATATTTCTTTTATG	1215
Qy	440	Tyr-----LysValSerIleValThrAspArgInserProProValSerPro	455
Db	1216	AACAACACTATGATGATNAGGTGGAATTTCTGACACTAAGAGTACTCAGATATAATA---	1272
Qy	456	IleGlnProHisPheIleIleAsnGlnIleGluLeuTyLeuAsnGlySerSerAsnAsn	475
Db	1273	-----TATGAATATGATATTTTTCGAAATAGACAGTGA	1305
Qy	476	ThirLeuIysTySerAlaGlyGlySerIeu---SerAsnTyrgInAsnThrPhePhe	494
Db	1306	GTATTTTGCATATTCATCCAAATTCACCAATGAAATAATATATTAAGAACGATCTTAT	1365
Qy	495	GlnPheProAArgIlyLysAspCysAsnLeuValIleAspProGlyCysSerProAsn	514
Db	1366	ATGATTCCTAAACCAACATCGAATAAT-----	1392
Qy	515	AsnAsnTySerSerHisIleLeuSerHisPheSerIeuPheThrTySerTyValIleGly	534
Db	1393	AAAGATATGTGTCATACTCTATCGATATTAACACTGATATATATATTTTCACTAGT	1452
Qy	535	LeuGlnLeuGlnIleLeuAspThrGlyValLeuGlyTyThrHisSerSerValAspArg	554
Db	1453	AGAGAAAGAAAGAGATT-----GCATTTAGTTGACACACTACGATGTTGATTTTC	1503
Qy	555	TyrAsnAlaIleSerAspLysIleIleThrMetIleProAlaIleIysGlyAsnAsnLeu	574
Db	1504	CAAAATACATATGATTTAGATTAACATCCCAAAATCCACGCTCTTAAGCTTTGAAGTA	1563
Qy	575	AspThrAsnSerIysValIleGlnGlyProGlyHisThrGlyIysAsnLeuValTyLeu	594
Db	1564	AGTTCTGATTCGAAATATGTGAAGGTCCTGGTCAACAGGTGAGACTTGGTATTTCTT	1623
Qy	595	GlnSerGlnIlyArgLeuGlnIleThrCysGlnThrProAsnSerThrGlnSerTyPhe	614
Db	1624	AAAGATAGTATGATTTTAAAGATTAAGATTT---TTAAAAAATGTTTTCGCAATATCA	1680
Qy	615	IleArgLeuArgTyArgAlaThrAsnGlyIleGlyAsnThrIleuProAsnIleSerIeuThr	634
Db	1681	GTAGCGATATCTTATGTCTACTTAATGCT-----CCAAAGACACAGTATTC	1725
Qy	635	IleProGlyValIleGlyIleProProGlnArgLeuAsnAsnThrPheSerGlyTyAsn	654
Db	1726	TTAACCGAATAGATCTATTAAGTGTGAG---CTCCCTAGTACCACTTCCCGCAAAAC	1782
Qy	655	TyrAsn-----AsnLeuGlnTyArgIlyAspPheGlyTyPheGlnPheProSerThrVal	672
Db	1783	CCAATGTCAAGATTTAATATATGATTTTGGATATGTAAATTTCCAAAGAACAGTT	1842
Qy	673	-----ThirLeuProLeuAsnArgAsnIleProPheIle	683
Db	1843	CCAAATAAACATTTGAAGAGAACACCTTATTTAATTAAGCC-----TTA	1887

```

Oy      684  pheasmnrglaaepvalseraamserileleuileleaplysilegileupellepro 703
Db      1888  TATGTACACCAACATCTTCCTATATAATA--TATATGACAAATCGAATTATTCCA 1944
Oy      704  lIetherserwethisglnasnarglulysglnlyseugluthrileglnthrls 723
Db      1945  ATCACTCATCTGTATTAGATTATACAGAGAGCAAAATATAGAAAAACACGAAATA 2004
Oy      724  lIeaanthrphethrAenhiSthLysThrleu 735
Db      2005  GTGAATGATTATTGTAAATTAACAAAGTCTT 2040

RESULT 5
US-08-315-468-3
; Sequence 3, Application US/08315468
; Patent No. 5554534
; GENERAL INFORMATION:
; APPLICANT: Michaels, Tracy Ellis
; APPLICANT: Michaela, Luis
; APPLICANT: Narva, Kenneth E.
; TITLE OF INVENTION: Process for Controlling Scarab Pests
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: David R. Saliwanchik
; STREET: 2421 N.W. 41st Street, Suite A-1
; CITY: Gainesville
; STATE: FL
; COUNTRY: USA
; ZIP: 32606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/315,468
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/014,941
; FILING DATE: 01 FEB 1993
; APPLICATION NUMBER: 07/828,430
; FILING DATE: 30-JAN-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/808,316
; FILING DATE: 16-DEC-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Saliwanchik, David R.
; REGISTRATION NUMBER: 31,794
; REFERENCE/DOCKET NUMBER: MA73, C2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 904-375-8100
; TELEFAX: 904-372-5800
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3507 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHEetical: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Bacillus thuringiensis
; STRAIN: kumamotoensis
; INDIVIDUAL ISOLATE: 50C
; IMMEDIATE SOURCE:
; LIBRARY: Lmbdagem-11(tm) library of L. Foncecrada
; CLONE: 50C(lb)
US-08-315-468-3
Alignment Scores:

```

Pred. No.:	6.55e-72	Length:	3507
Score:	775.50	Matches:	217
Percent Similarity:	46.44%	Conservative:	136
Best Local Similarity:	28.54	Mismatches:	289
Query Match:	20.04	Indels:	119
DB:	2	Gaps:	27

OY	5	AsnAspAsnAsnGluThrGluIleIleAspSerHisThrSerProGlyPheProAsnArg	24
Db	10	AAATATCAAAATGAATATGAATATTAAGATGACACCTTCT	57
OY	25	AsnSerAsnAspSer---ArgTyrProTyrThrAsnAsnProAsnGlnProLeuGlnAsn	43
Db	58	GATACCAATGATTCTTAACAGATACCTCTTTGGCAATGAGCGCAACAAATGGCGTACAAAAT	117
OY	44	ThrAsnTyrGluGluTyrLeuAsnMetCysGlnGlyAsn---ThrGlnTyrGlyAspAsn	62
Db	118	ATGATATTATTAAGATTATTTAAAAATAGTCGGCGGAAATGTTGTGAATCCCTGGTTCA	177
OY	63	PheGluThrPheAlaSerAlaAspThrIleAlaIleValSerAlaGlyThrIleValSer	82
Db	178	CCTGAGGATATTCTTAAGCGACGACAGAT-----GCGATTAAAGCCCGCAATTGATATATGA	231
OY	83	GlyThrLeuLeuAlaGlyIleGlyGlyLeuThrSerIleSerGlyProIleGlyIleIle	102
Db	232	GGTAAATTACTTAACAGGTTTAAGG-----GTTCCATTGTT	267
OY	103	GlyAlaIleIleIleSerPheGlyThrLeuIleThrValPheThrProAlaGlyGluGln	122
Db	268	GCGCGCATGTGATGATCTTTATCTCACTTAATGATATCTGGCGCTTCA-----AAA	321
OY	123	AspTyrThrValTyrThrGlnPheIleTyrMetGlyGluIlePheValAspThrProLeu	142
Db	322	CAAAAGAGCAATGGGAAATTTTATGAGCAACATGAAGAACTCAATTAATCAAAAAATA	381
OY	143	ThrGluSerIleLeuGlnLeuTyrLeuGlnThrLeuGlnGlyPheArgGlnIleLeuGln	162
Db	382	CGAGAAATATGCAGAGAAATAAGCCCTTTCGGATTTGGAAAGCGTTAGGAAATATTACAA	441
OY	163	SerTyrAsnThrAlaLeuAspAspThrPargTyrLeuTyrArgLeuGlnAlaProGlyLeu	182
Db	442	TTATATCTTAACGCGCTTGAAGAGTGGAAAGAA-----	474
OY	183	ProProSerSerAlaLeuGlnGlnAlaAlaLeuThrLeuTyrIleArgPheGluAsnVal	202
Db	475	AATCCAAATGGTTCA-----AGAGCTTACGAGATGTTTGGAAATGATTAAGATCTG	528
OY	203	HisAsnAspPheIleArgGluIleProGlyPheGlnLeuGlnIleThrTyrXlySerThrLeu	222
Db	529	GATAGTTATTATTCGCAATATATGCCATCTTTCCAGTGCACAAATTTTGAAGTACCAATC	588
OY	223	LeuProIleTyrAlaGlnAlaAlaAsnPheHisLeuAsnLeuLeuGlnGlnGlyAlaGln	242
Db	589	CTTACAGTATATACAAATGCGACGACAAACCTACATTCTTATTAAGGACGACATCAAT	648
OY	243	LeuAlaAspGluTyrAsnAlaAspIleHisProSerGlnIleGluProAsnAlaGlyThr	263
Db	649	TTTGGAGGAATATGGGGATTTG-----TTCACAGCACT	681
OY	263	SerAspAspTyrTyrTyrLeuLeuTyrGluAsnIleProTyrTyrSerAsnTyrCysAla	282
Db	682	ATTATATTAATCTATATATGCTCAATAGAAACTTACTGCGAGAAATATTTGACCACTGTGA	741
OY	283	AsnThrTyrArgThrGlyLeuTyrAsnLeuArgPoliProAsnMetTyrTrpSerIle	302
Db	742	AAGTGGTATGAACCTGTTTAGCAAAATTAATAAGCTCGACGCCCTAAACAATGGAATTGAC	801
OY	303	PheAsnAspTyrArgArgTyrMetThrIleThrValLeuAspThrIleSerGlnPheSer	322
Db	802	TATTAACCAATTCGCTGAGAAAGAAAGACATTTGACGCTGTGTGACCTGTGGATTAATTTTCA	861

QY	323	LeuYrAspIleuYrAsgYrYrArgAspSerIleGlyGlyIleGluValIleGlyIleYs	342
Db	862	AACTATGATACCGCTAGCTAT-----CCACTGGCAACACA	897
QY	343	AsnGluLeuThrArgGluIleYrThrArgGluIleAsnPheAspArgLeuProGlnLeu	362
Db	898	GCTCAGCTTACAGGAGATATATACATCCACTGGCGCGGTAGATGTGCTAATAT	957
QY	363	ArgValGlnProAsnLeuAlaThrMetGluYrAsnLeuThrArgAlaSerPheYs---	381
Db	958	GGCTCTGTGTATGACAAAGACCTTCTTCTCAGAAATATGAATAAAGCGGCTTGTGTCA	1011
QY	382	-----LeuPheSerPheLeuGlnGlnPheIlePheYrThrGluAsnThrAsnPe---	398
Db	1018	CCTCATGTGTTGATTATATATACGGGACTCACAGTTTATACAAAATAAGCTTCACT	107
QY	399	-----GlyAsnArgLeuValGlyIleSerAsnArgAsp	409
Db	1078	TCGTATCGTTATATGAGATATTTGGCTGTATCA-----ATAAGCTATAGCAT	112
QY	410	AlaProThrYrSerAsnThrIleThrIuThrLeuYrGlyArgThrGlySerPro	429
Db	1129	ATCGGTACG---AGTATGACCTTTACACAG---ATGATATGAACCATCAAAATTTACA	118
QY	430	ThrThrIlyThrIleArgProPheGluSerYrIlyVal-----	442
Db	1183	AGTACTACCAATTTT---GATTATACGATATTAACGATATTTCACAGACTTATCAATGT	123
QY	443	SerIleValThrAspArgGlnSerProProValSer-----ProIle	456
Db	1240	GCAAGTACCTTGATATAGTTTACCTGTATACGATATACCTTTTGGATGCCAGAA	129
QY	457	GlnProHisPheIleIleAsnGlnIleGluLeuYrLeuAsnGlySerSerAsnThr	476
Db	1300	ACCGAGTTTTTATGTAAATCAATATG-----AATATATACC	133
QY	477	LeuYrYrYrSerAlaGlyGlySerLeuSerAsnYrGlnAsnThrThrPheGlnPhe	496
Db	1336	AGAAAG-----ACGTTAAGCTTAAACACGCTTCCAAAGATATATTA	137
QY	497	ProArgGlyIlyAspCysAsnLeuValIleAspProGlyCysSer-----ProAsnPe	514
Db	1378	GATCGACACAGAGATTCGGAATTAAGATTCCTCCAGAACTTCAGGTCAACCAATTA	143
QY	515	AsnAsnYrSerHisIleLeuSerHis-----PheSerLeuPheThrYrSer	530
Db	1438	GAGTCATATACCATATAGATTTAGTCATATTAATTAATTAATTCAGTTCAATACGAC	149
QY	531	TyrValIleGlyLeuGlnLeuGlnIleLeuAspThrGlyValLeuGlyYrThrHisSer	550
Db	1498	TATGTA-----CCTGATATTTCTTGACACATCCG	152
QY	551	SerValAspArgYrAsnAlaIleSerAspYsIleIleThrMetIleProAlaIleYs	570
Db	1558	AGTCGAGATCTTAACAATACAGTTAAAGTGGCAAAATACCCAAATATACAGGGGCAAG	158
QY	571	GlyAsnAsnLeuAspThrAsnSerIlyValIleGluGlyProGlyHisThrGlyIlyAsn	590
Db	1588	TCTAGCACCATATGCGAAGATATCTATATATATATAAAGCGCGGTATATACAGGGGAGAC	164
QY	591	LeuValYrLeuGlnSerGln---GlyArgLeuGluIleThrCysGlnIuThrProAsnSer	609
Db	1648	TTAGTGGCTTTAAACGACCGCATCGGAAGTTGTGAGTTTCAGATGATCTTTCAGAGTCT	170
QY	610	ThrGlnSerYrPheIleArgLeuArgYrYrAlaThrAsnGlyValaGlyAsnThrLeuPro	629
Db	1708	---CAACATTCCTCGATTCGATTCGATTCGCTTCTTAAGAACTAGT-----	175
QY	630	AsnIleSerLeuThrIleProGlyValIleGlyIleProProGlnArgLeuAsnAsnThr	649
Db	1753	-----TATATTTAGTTTATACGACCTAAACCAAAGCGAAGCTTAAATTCACACGACAC	180
QY	650	PheSerGlyThrAsnYrAsnAsnLeuGlnYrGlyAspPheGlyYrThrGlnPhePro	669







```

/ ORIGINAL SOURCE:
/ ORGANISM: Bacillus thuringiensis
/ STRAIN: colworthi
/ INDIVIDUAL ISOLATE: 43f
/ IMMEDIATE SOURCE:
/ CLONE: E. coli XL1-Blue (pM1, 98-4), NRRL B-16291
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: 1..1953
US-07-941-650A-3

Alignment Scores:
Pred. No.: 1,05e-61 Length: 1953
Score: 676.00 Matches: 209
Percent Similarity: 43.8% Conservative: 117
Best Local Similarity: 28.1% Mismatches: 283
Query Match: 17.5% Indels: 136
DB: Gaps: 29

US-10-783-417-2 (1-735) x US-07-941-650A-3 (1-1953)

QY 1 MetAenGlnAsnAspAsnAsnGluTyrGluIleIleAspSerHisThrSerProTyr 20
DB 1 ATGAATCCAAACAAAT---CGAAGTGAATATGATACGATTAAGGTTACACCTTAACACTGA 57
QY 21 PheProAsnArgAsnSerAsnAspSerArgTyrProTyrThrAsnAsnProAsnGlnPro 40
DB 58 TTGGCC-----ACTAACCATATATCAATATCTTTAGCTGACATCCAAATTCGACA 108
QY 41 LeuGlnAsnThrAsnTyrLysGluTyrPheAsnMetCysGlnGlyAsnThrGlnTyrGly 60
DB 109 CTAGAAAGATTAATATTAAGATTTTAAAGATGACGACAGACAACTTCACG----- 162
QY 61 AspAsnPheGluThrPheAlaSerAlaAspThrIleAlaValaSerAlaGlyThrIle 80
DB 163 -----GAAGTGTAGACAGCTCTACAGTAAAGATGAGAGCTGGACAGGAATTTCT 213
QY 81 ValSerGlyThrLeuLeuAlaGlyIleGlyGlyLeuThrSerIleSerGlyProIleGly 100
DB 214 GTTGTAGGACAGATTTAGGCTGTGAGG-----GTTCAC 249
QY 101 IleIleGlyAlaIleIleIleSerPheGlyThrLeuIleThrValPheTyrProAlaGly 120
DB 250 TTTCGTGGGGCCCTCATCTTATTCATCATCTTTCTTAACGCTATATGCCCAAT--- 306
QY 121 GlnGlnAspLysThrValTyrThrGlnPheIleLysMetGlyGluIlePheValaAspThr 140
DB 307 -----GATGCTGACCCATGGAAGGCTTTATGGCACAAGTGAAGTACATGATAGTAAG 360
QY 141 ProLeuThrGlnSerIleLysGlnLeuLysLeuGlnThrLeuGlnGlyPheArgGlnIle 160
DB 361 AAAATGAGAGAGTACTAAAGTAAAGCTCTTGACAGGTACTACAGGCTTCCAAATTAAT 420
QY 161 LeuGlnSerTyrAsnThrAlaLeuAspAspTyrPargLysLeuLysArgLeuGlnAlaPro 180
DB 421 TTTCGAAGATTAATGTAATGCGTTGATCTCTGGAAGAAA----- 459
QY 181 GlyLeuProProSerSerAlaLeuGlnGlnAlaAlaLeuThrLeuLysIleArgPheGlu 200
DB 460 ---GGCCCTGTAATATTACGAAGTGAAGAGCCAGATCGAATAGAGAACTTTTCTTC 516
QY 201 AsnValHisAsnAspPheIleArgGluIleProGlyPheGlnLeuGluThrTyrLysThr 220
DB 517 CAAGCAGAAAGCCATTTCGTAAATTCATGCGCTCATTTGCGTTTCCAAATTCGAAGTT 576
QY 221 LeuLeuLeuProIleTyrAlaGlnAlaAlaAsnPheIleLeuAsnLeuLeuGlnGlnGly 240
DB 577 CTGTTTCTACCAACATATGCAACAAGCTGCAAAATACCATTTATGTGTAATTAAGATGCT 636
QY 241 AlaGlnLeuAlaAspGluTyrAsnAlaAspIleHisProSerGlnIleGluProAsnAla 260
DB 637 CAAGTTTTCGAGAAAGATGCGGATAT----- 663

```

```

QY 261 GlyThrSerAspAsp-----TyrTyrLysLeuLeuLysGluAsnIleProLysTyr 277
DB 664 ---TCCTGAGAGATATATGCTGAATTTATCAAGACAAATTAACCTTACGCAACAATAC 720
QY 278 SerAsnTyrCysAlaAsnThrTyrArgThrGlyLeuLysAsnLeuArgAspGluProAsn 297
DB 721 ACTGACCAATGTGTCAATGTGATATATGTGATTAATTAAGATTAAGAGGTTCAACTTAT 780
QY 298 MetLysTyrSerIlePheAsnAspTyrArgArgTyrMetThrIleThrValaLeuAspThr 317
DB 781 GATGATGGGCTAAATTTTAACCGTTTTCGACAGAAATGACATTAATCTATATAGATCTTA 840
QY 318 IleSerGlnPheSerLeuTyrAspIleLysArgTyrArgAspSerIleGlyGlyIleGlu 337
DB 841 ATTGATATTAATCCATTTATATGATGTGGTATATCTCA----- 879
QY 338 ValLysGlyIleLysAsnGluLeuThrArgGluIleTyrThrThrGluIleAsnPheAsp 357
DB 880 ---AAAGAGTTTAAACAGACCTAACAGACATTTTACAGATCCAAATTTTACCA 933
QY 358 ArgLeuProGlnLeuArgValGlnProAsnLeuAlaThrMetGluTyrAsnLeuThrArg 377
DB 934 CTCGAATGCTCTTCAAGATGAGCAACTTTTTCGAGATAGAAAACCTATTCGAAAA 993
QY 378 AlaSerPheLysLeuPheSerPheLeuGlnGlnPheIlePheTyrThrGluAsnThrAsn 397
DB 994 CCT-----CATTTATGATATATTTGCTGGATGTAATTCATACG----- 1035
QY 398 PheGlyAsnArgLeu---ValGlyIleSerAsnArgAspAlaProThrTyr----- 413
DB 1036 -----CGTCTTCGACCTGGTTACTCTGGAAAAGATTCCTTCATATATGTCGTCT 1086
QY 414 -----SerAsnThrIleThrGluThrLeuTyr 422
DB 1087 AATATATGAGAACTAGACCTAGATATGAGATCTAATGATACATATCACTTCCCATTTAT 1146
QY 423 GlyLysArgThrGlySerProThrThrTyrLysThrIleArgProPheGlnSer----- 439
DB 1147 GGAAGTAAATCATTTGAACTTATACAAAGCTA-----AGCTTGAATGACAAAAGTT 1200
QY 440 TyrLysValSerIleValThrAspArgGlnSerProProValSerProIleGlnProHis 459
DB 1201 TATCAACATATGATATATACAGATGAGCGCTTTCCGATGCGCAAGATA-----TAT 1254
QY 460 PheIleIleAsnGlnIleGluLeu---TyrLeuAsnGlySerSerAsnAsnThrLeuLys 478
DB 1255 TTTCGTGTTCGAAAAGTTGATTTTATGTCATATATGATCAAAAATGAAACT----- 1308
QY 479 TyrSerAlaGlyLysSerLeuSerAsnTyrGlnAsnThrThrPhePheGlnPheProArg 498
DB 1309 -----AGTACACAAACATATGATTTCAAAAAGTATACATGCTATTTAGGT 1353
QY 499 LysLysAspCysAsnLeuValIleAspProGlyCysSer-----ProAsnPheAsnAsn 516
DB 1354 GCACAGAGATTCATGACCAATTAACCAAGAAACAGAGAAACCACTTGAAAAAGCA 1413
QY 517 TyrSerHisIleLeuSerHisPheSerLeuPheThrTyrSerTyrValIleGlyLeuGln 536
DB 1414 TATAGTCATCAGCTTAATTAATCGCAGATGTTCTTA-----ATGCGAG 1455
QY 537 LeuGlnIleLeuAspThrGlyValaLeuGlyTyrThrHisSerSerValaAspArgTyrAsn 556
DB 1456 GACCGTCGTGAAACAATTCATTTTACTTGGACATAGAAAGGTAGACTTTTATAT 1515
QY 557 AlaIleSerAspLysIleIleThrMetIleProAlaIleLysGlyAsnAsnLeuAspThr 576
DB 1516 ACAATTTGATGCGTGAATAATTAATCTCACTTCAGATGAGAAAGCAATGCTGTCTCA 1575
QY 577 AsnSerLysValIleGluGlyProGlyHisThrGlyGlyAsnLeuValTyrLeuGln---- 595
DB 1576 GCGCTTCATTAATTAAGGTTCAGATTTCAAGAGGAAATTTTACTATCTTAAGAA 1635
QY -----SerGlnGlyArgLeuGluIleThrCysGluThrPheAsnSerThrGlnSer 612

```

Db 1636 TCTAGTAATTCATGCTAAATTTAAAGTTACCTTAATTAATTCAGCAGCCTTGTTACACGA 1695  
Qy 613 TyrPheIleArgIleuArgTyrAlaThrAenGlyAlaGlyAenThrLeuProAenIleSer 632  
Db 1696 TATCCGCTAGAGAAATACGCTATGCTTCAAC-----ACTTACCTTACGA 1737  
Qy 633 LeuThrIleProGlyValIleGlyIleProProGlnArgLeuAenAenThrPheSerGly 652  
Db 1738 CTTTTCGTG-----CAAAATTCAAACAATGATTTCTGTGC 1773  
Qy 653 ThrAenTyrAen-----AenLeuGlnTyrGlyAenPheGlyTyr-----PheGln 667  
Db 1774 ATCTACATTAAATAAATACTATGATAATAGAT---GGTGAATTAAACATATCAAAATTTGAT 1830  
Qy 668 PheProSerThrValThrLeuProLeuAenArgAenIleProPheIlePheAenArgAla 687  
Db 1831 TTCCGAACTAGT-----AATTCTAATATGCGATTTCTGCGTATCAAAAT 1875  
Qy 688 Asp-----ValSerAenSerIleLeuIleIleAenIleAenIle 699  
Db 1876 GACCTTATATATAGAGACGAATCTTTCGTTCTAATGAAAAAATCTATATAGATAGATA 1935  
Qy 700 GluPheIleProIle 704  
Db 1936 GAATTTATCCACGTA 1950

RESULT 8  
PCT-US92-00040-1  
Sequence 1, Application PC/TUS9200040  
GENERAL INFORMATION:  
APPLICANT: Donovan, William P.  
APPLICANT: Ruper, Mark J.  
APPLICANT: Stanley, Annette C.  
TITLE OF INVENTION: BACILLUS THURINGIENSIS CYTILIC (b) TOXIN  
TITLE OF INVENTION: GENE AND PROTEIN TOXIC TO COLEOPTERAN INSECTS  
NUMBER OF SEQUENCES: 2  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Panitch Schwarze Jacobs & Nadel c/o A.S.  
STREET: 1601 Market Street, 36th Floor  
CITY: Philadelphia  
STATE: Pennsylvania  
COUNTRY: U.S.A.  
ZIP: 19103  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US92/00040  
FILING DATE: 19920103  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/649,562  
FILING DATE: 31-JAN-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Egolf, Christopher  
REGISTRATION NUMBER: 27633  
REFERENCE/DOCKET NUMBER: 7205-29 P1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 215-757-1590  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2430 base pairs  
TYPE: NUCLEIC ACID  
STRANDEDNESS: double  
TOPOLOGY: circular  
MOLECULE TYPE: DNA (genomic)  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 144..2099

PCT-US92-00040-1  
Alignment Scores:  
Pred. No.: 3,24e-60 Length: 2430  
Score: 663.50 Matches: 217  
Percent Similarity: 41.9% Conservative: 115  
Best Local Similarity: 27.4% Mismatches: 300  
Query Match: 17.1% Indels: 161  
DB: 6 Gaps: 31

US-10-783-417-2 (1-735) x PCT-US92-00040-1 (1-2430)

Qy 1 MetAenGlnAenAenAenAenAenGluTyrGluIleIleAenSerHisThrSerProTyr 20  
Db 144 ATGAATCCAAACAAT---CGAAGTCAGACATATACGATAAAGTTACCACTTCAAGTGA 200  
Qy 21 PheProAenArgAenSerAenAenSerArgTyrProTyrThrAenAenProAenGlnPro 40  
Db 201 TTGCCA-----ACTTACCATATATCAATATCTTTAGCTGACATCAATTCGACACA 251  
Qy 41 LeuGlnAenThrAenTyrGlyGluTyrLeuAenMetCysGln---GlyAenThrGlnTyr 59  
Db 252 CTAGAGAATTTAAATTAATTAAGAAATTTTAAGATGACTGAGAAGACAGTTTACGGAAGT 311  
Qy 60 GlyAenAenPheGluThrPheAlaSerAlaAenThrIleAla---AlaValSerAlaGly 78  
Db 312 CTAGACAAC-----TCTACAGTAAAGAAGATGCGACAGCA 350  
Qy 79 ThrIleValSerGlyThrLeuLeuAlaGlyIleGlyGlyLeuThrSerIleSerGlyPro 98  
Db 351 ATTTCTGTTTGAAGGACAGATTTTAGGTGTATAGA----- 386  
Qy 99 IleGlyIleIleGlyAlaIleIleIleSerPheGlyThrLeuIleThrValPheTyrPro 118  
Db 387 GTTCATTGCTGGGGCACTCACTTCATTATATCAATCATTTCTTGAACATATAGGCA 446  
Qy 119 AlaGlyGlnAenPlyThrValTyrThrGlnPheIleIleMetGlyGluIlePheVal 138  
Db 447 AGT-----GATGCTGACCCATGGAAGGCTTTATGAGCAACAATGGAAGTATGATA 497  
Qy 139 AspThrProLeuThrGlnSerIleGlyGlnLeuGlnThrGlnThrLeuGlnGlyPheArg 158  
Db 498 GATAGAAAATGAGAGATATGCTAAAGCTTTGCAAGATTCACAGGCTTTCAA 557  
Qy 159 GlnIleLeuGlnSerTyrAenThrAlaLeuAenAenTyrPargIleLeuArgLeuGln 178  
Db 558 AATAATTTGCAAGATATATATGCTTAATTCCTGGAAGAA----- 602  
Qy 179 AlaProGlyLeuProProSerSerAlaLeuGlnGlnAlaIleLeuThrLeuIleArg 198  
Db 603 -----ACACCTTTAAGTTTGGCAAGTAAAGAACCCAAAGTCGAATAGGGAACCT 653  
Qy 199 PheGlnAenValHisAenAenPheIleArgGluIleProGlyPheGlnLeuGlnThrTyr 218  
Db 654 TTTTCTCAAGCAGAAAGCATTTTCTGTAATTCACCGCTGATTTGCAAGTTCCAAATTC 713  
Qy 219 LysThrLeuLeuLeuProIleTyrAlaGlnAlaIleAenPheHisLeuAenLeuGln 238  
Db 714 GAAGTGTCTTTCTTACCAACATATGACACAGCTGAAATACACATTTATGCTATTAATA 773  
Qy 239 GlnGlyAlaGluLeuAlaAenProIleTyrAenAlaAenIleHisProSerGlnIleGluPro 258  
Db 774 GATGCTCAAGTTTTCGAGAAAGATGGGATATTC----- 809  
Qy 259 AenAlaGlyThrSerAenAenPyrTyrIleLeuLeuGlyGluAenIleProLysTyrSer 278  
Db 810 TCAGAAAGATGTTGCTGAATTTATCATATGACAAATTAATAACTTACG---CAACAATACACT 866  
Qy 279 AenTyrCysAlaAenThrTyrArgThrGlyLeuLeuAenLeuArgAenGluProAenMet 298  
Db 867 GACCATTTGTCAATTTGATATAGTTGGATTTAAATGTTTAAAGGTTCAACTATAGAT 926  
Qy 299 LysTyrSerIlePheAenAenPyrTyrArgTyrMetThrIleThrValLeuAenThrIle 318

```

Db      927 GCATGGGTCAAAATTAAACCGTTTTCGACAGAAATACCTTAAGTATTAATTAAT 986
Qy      319 SerGlnPheSerLeuThrAspIleValArgIleValArgAspSerIleGlyIleVal 338
Db      987 GATCTTTCCTCAATTATTAATGATGTTTCGTTACTCA-----1022
Qy      339 LysGlyIleLeuAsnGluLeuThrArgGluIleValThrThrGluIleAsnPheAspArg 358
Db      1023 AAGAGTGTAAACAGAACTTACAGAGACATTTTTCGATCCAAAT---TTTTCATC 1079
Qy      359 LeuProGlnLeuArgValGlnProAsnLeuValIleThrMetGluThrAsnLeuThrArgVal 378
Db      1080 AATACCTCTCAGAGAGATAGACCACTTTTTCAGATAGAGAAACCTATTCGAAAAACCT 1139
Qy      379 SerPheLeuLeuPheSerPheLeuGluGlnPheIlePheThrThrGluAsnThrAsnPhe 398
Db      1140 -----CATTTATTTGATTTATTAACAGGATTTGAAATTTTCATACG-----1178
Qy      399 GlyAsnArgLeu---ValGlyIleSerAsnArgAspAlaProThrTyr-----413
Db      1179 -----CGTCTTCACACCTGCTTACCTGGAGAAATCTTCAATTAATGCTGCTGAAT 1232
Qy      414 -----SerAsnThrIleThrGluThrLeuThrLeuThrGly 423
Db      1233 TATGTAGAAACTAGACCTAGTATAGATCTAGTAAGACATTAATCTCCCATTTATAGA 1292
Qy      424 GluValGlnThrGlySerProThrThrIleValArgProPheGluSer-----Tyr 440
Db      1293 GATTAATCTTACGACCTGATACCAAAAGTTA-----AGCTTGATGACAAAAAGTTTAT 1346
Qy      441 LysValSerIleValThrAspArgGlnSerProProValSerProIleGlnProHisPhe 460
Db      1347 CGAATATAGCTAATATACAGACGTAGCGGCTTGCCCAATGCGCAAGATA-----TATTTT 1400
Qy      461 IleIleAsnGlnIleGluLeu-----TyrLeuAsnGlySerSerAsnAsn 475
Db      1401 GGTGTTCAGAAAAGTTTATTTTACTCAATATGATGATCAAAAAAATGAAACTAGTACACA 1460
Qy      476 ThrLeuLeuTyrSerIleGlySerLeuSerAsnTyrGlnAsnThrThrPheGln 495
Db      1461 ACATATGATTCAAAAAAGAAACAATGCCATGTAGTGACACAGATCTTATTCACCAATTA 1520
Qy      496 PheProArgIleValAspCysAsnLeuValIleAspProGlyCysSerProAsnPheAsn 515
Db      1521 CCACCAAGAAACAACAT-----GAAACCACTTGAAAA 1553
Qy      516 AsnTyrSerHisIleLeuSerHisPheSerLeuPheThrTyrSerTyrValIleGlyLeu 535
Db      1554 GCATATAGTCATCAGCTTAATTAACGGGATGTTCTTA-----ATG 1595
Qy      536 GlnLeuGlnIleLeuAspThrArgIleValLeuGlyTyrThrHisSerSerValAspArgTyr 555
Db      1596 CAGGACCGTCGAGAACAAATTCATTTTACTTGACACATGTGAAGTGAAGCTTTT 1655
Qy      556 AsnAlaIleSerAspIleIleIleThrMetIleProAlaIleGlySerAsnLeuAsp 575
Db      1656 AATACAAATGATGCTGAAAGATTAATCTCAACTTCAGTAGTAGAACATATGCTTGTCT 1715
Qy      576 ThrAsnSerLeuValIleGluGlyProGlyHisIleThrGlyValAsnLeuValTyrLeuGln 595
Db      1716 TCAAGTCTTCATTTATTAAGTCCAGATTCACAGAGAAATTTACTATTCCTTAAAA 1775
Qy      596 -----SerGlnGlyArgLeuGluIleThrCysGluThrProAsnSerThrGln 611
Db      1776 GAATAGTAATATCATATGCTAATTAATTAAGTAACTTAATTAATCAGACGCTTGATTA 1835
Qy      612 SerTyrPheIleArgLeuArgTyrAlaThrAsnGlyAlaGlyAsnThrLeuProAsnIle 631
Db      1836 CGATATCGTGTAGAGATAGCTATGCTTACAC-----ACTAATCTTA 1877
Qy      632 SerLeuThrIleProGlyValIleGlyIleProGlnArgLeuAsnAsnThrPheSer 651

```

```

Db      1878 CGACTTTTGTG-----CAAAATTCAAACATGATTTTAT 1913
Qy      652 GlyThrAsnTyrAsn-----AsnLeuGlnTyrGlyAspPheGlyTyrPheGlnPhe 668
Db      1914 GTCATCTACATTAATTAATAAATATATGATATAGAT---GATGATTTAATATCAAAACATTT 1970
Qy      669 ProSerThrValThrLeuProLeuAsnArgAsnIleProPhe-----682
Db      1971 GATCTCGCACTACT-----AATCTATATATGCGGTTCTCGGTCATACGATGAA 2021
Qy      683 -----IlePheAsnArgAlaAspValSerAsnSerIleLeuIleIleAspIleGlu 700
Db      2022 CTATATATATGAGACAGAAATCTTTCGTTCTAATGAAAAAATCTATATATGATTAAGATGAA 2081
Qy      701 PheIleProIle-----704
Db      2082 TTTATTCCTCAGTCAATTTGTAAGAGATTTGAAATGTAGGCGATGTCAAAAATGAAAGA 2141
Qy      705 -----ThrSerMetHisGlnAsnArgGluIys 714
Db      2142 ATAGGAGGTGAATTTGATGCTTAGGAAAGATCTTTTAAGAAAAAGCAATGGAAGG 2201
Qy      715 GlnIysLeuGluThrIleGlnThrIys---IleAsnThr 726
Db      2202 TATACAGTACAAATATTAAGAAATTAATTAACACA 2240

RESULT 9
US-08-996-441B-37
; Sequence 37, Application US/08996441B
; Patent No. 6023013
; GENERAL INFORMATION:
; APPLICANT: English, Leigh H.
; APPLICANT: Bruseock, Susan M.
; APPLICANT: Malvar, Thomas M.
; APPLICANT: Bryson, James W.
; APPLICANT: Kulesza, Caroline A.
; APPLICANT: Walters, Frederick S.
; APPLICANT: Staltn, Stephen L.
; APPLICANT: Von Tersehn, Michael A.
; APPLICANT: Romano, Charles
; TITLE OF INVENTION: INSECT-RESISTANT TRANSGENIC PLANTS
; NUMBER OF SEQUENCES: 113
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Arnold, White & Durkee
; STREET: P.O. Box 4433
; CITY: Houston
; STATE: Texas
; COUNTRY: USA
; ZIP: 77210
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/996,441B
; FILING DATE: 18-DEC-1997
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Kitchell, Barbara S.
; REGISTRATION NUMBER: 33,928
; REFERENCE/DOCKET NUMBER: MECO:151
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 512/418-3000
; TELEFAX: 512/474-7577
; INFORMATION FOR SEQ ID NO: 37:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1959 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; FEATURE:
; NAME/KEY: CDS

```

LOCATION: 1..1956  
US-08-996-441B-37

Alignment Scores:  
Pred. No.: 2,586-60 Length: 1959  
Score: 663.00 Matches: 208  
Percent Similarity: 43.5% Conservative: 114  
Best Local Similarity: 28.1% Mismatches: 292  
Query Match: 17.1% Indels: 126  
DB: 3 Gaps: 27

US-10-783-417-2 (1-735) x US-08-996-441B-37 (1-1959)

```

QY      1 MetAenglnAsnAspAsnAsnGluTyrGluIleIleAspSerHisThrSerProTyr 20
DB      1 ATGAATCCAAACAAAT---CGAAGTGAAACATGAT-----ACGATAAAGGT 42
QY      21 PheProAsnArg-----AsnSerAsnAspSerArgTyrProTyrThrAsnAsnProAsn 38
DB      43 ACACCTTAACAGTGAATTGCAACCTAACATATCATATCTTTAGCTGACATCCAAAT 102
QY      39 GluProLeuGlnAsnThrAsnTyrIleGluTyrLeuAsnMetCysGln---GlyAsnThr 57
DB      103 TCAACACTRGAAGAATTAAATTATTAAGAAATTTTAAGATGACTGAAACAGCTTCTACG 162
QY      58 GlnTyrGlyAspAsnPheGluThrPheAlaSerAlaAspThrIleAla---AlaValSer 76
DB      163 GAAGTGCTAGACAC-----TCTACAGTAAAGATGACAGTTGGG 201
QY      77 AlaGlyThrIleValIleSerGlyThrLeuLeuAlaGlyIleGlyLeuThrSerIleSer 96
DB      202 ACAGAAATTTCTGTTGTAGGCGAGATTTAGGCTGTAGGA----- 243
QY      97 GlyProIleGlyIleIleGlyAlaIleIleIleSerPheGlyThrLeuIleThrValPhe 116
DB      244 -----GTTCCATTGCTGGGCACTCACTTCAATTTTATCAATCTTTCTTAACCTATA 297
QY      117 TrpProAlaGlyGluGlnAsnAspIleThrValIleTrpThrGlnPheIleLysMetGlyGluIle 136
DB      298 TGGCCCAAGT-----GATGCTGACCCAGGAAAGGCTTTTATGCGACAACTTGAAGA 348
QY      137 PheValAspThrProLeuThrGluSerIleLysGlnLeuLeuGlnThrLeuGluGly 156
DB      349 CTGAATGATTAAGAAATATAGAGATATGCTTAAAGTAAAGCTCTTGCAAGATTTACAGG 408
QY      157 PheArgGlnIleLeuGlnSerTyrAsnThrAlaLeuAspAspTrpArgLysLeuLysArg 176
DB      409 CTTCAAAATTAATTTGGAAGATTAATGCTTAATGCGTTAAATCTCTGAAAGAA----- 459
QY      177 LeuGlnAlaProGlyLeuProPheSerSerAlaLeuGlnGlnAlaAlaLeuThrLeuLys 196
DB      460 -----ACACCTTAAGTTTGGCAAGTAAAGAAAGCAAGCAAGATGCAATTAAGG 504
QY      197 IleArgPheGluAsnValHisAsnAspPheIleArgGluIleProGlyPheGlnLeuGlu 216
DB      505 GAACCTTTTCTCAAGCAGAAAGTCAATTTTCTTAATTCATCCCGCTCATTTTCCAGATTTCC 564
QY      217 ThrTyrLysThrLeuLeuLeuProIleTyrAlaGlnAlaAlaAsnPheHisLeuAsnLeu 236
DB      565 AAATTCGAAGTCTGTTCTTACCAACATATGCAAGAGTCAAAATACACATTTATTTGCTCA 624
QY      237 LeuGlnGlnGlyAlaGluLeuAlaAspGluTyrAsnAlaAspIleHisProSerGlnIle 256
DB      625 TTAANAAGATGCTCAAGTTTGGAGAAGATGGGGAATTC----- 666
QY      257 GluProAsnAlaGlyThrSerAspAspTyrTyrLysLeuLeuLysGluAsnIleProLys 276
DB      667 -----TCAGAAGATGTTGCTGAATTTTATCATAGCAATTAATAACTTACA---CAACA 717
QY      277 TyrSerAsnTyrCysAlaAsnThrTyrArgThrGlyLeuLysAsnLeuArgAspGluPro 296
DB      718 TACACTGACCATTTGTTAATTGTTAATGTTGAATTAATGTTTAAAGAGTTCAACT 777

```

```

QY      297 AsnMetLysTrpSerIlePheAsnAspTyrArgArgTyrMetThrIleThrValLeuAsp 316
DB      778 TATGATGATGGGTCAAAATTAAACCGTTTGCAGAGAAATGACTTAACTGTATTAAGT 837
QY      317 ThrIleSerGlnPheSerLeuTyrAspIleLysArgTyrArgAspSerIleGlyGlyIle 336
DB      838 CTAATTTGACTTTTCCATTTTAATGATATTCGGTTATCTCA----- 879
QY      337 GluValLysGlyIleLysAsnGluLeuThrArgGluIleTyrThrThrGluIleAsnPhe 356
DB      880 -----AAAGCGTTAAACAGAACTAACAGAGCATTTTACGATTCATATTTTCT 933
QY      357 AspArgLeuProGlnLeuArgValGlnProAsnLeuAlaThrMetGluTyrAsnLeuThr 376
DB      934 CTTAGAGCACCA---CTTGGCGTACCGAACCACTTTTGTAGATATGAATAACTATTCGA 990
QY      377 ArgAlaSerPheLysLeuPheSerPheLeuGlnGlnPheIlePheTyrThr 393
DB      991 AAACCT-----CATTTATTTGATTATTATACAGGGGATTTGAATTTCAATACCGCTTCA 1044
QY      394 -----GluAsnThrAsnPhe-----GlyAsnArgLeuValGlyIle 405
DB      1045 CTTGTTACTTTGGGAAAGATTTCTTCAATTAATTCGTCTGGTAATTA-----GTA 1095
QY      406 SerAsnArgAspAlaProThrTyrSerAsnThrIleThrGluThrLeuTyrGlyLysArg 425
DB      1096 GAACCTAGACCTAGATATGAGATCTATGAACAATTACTCCCATTTTATGAGATTA 1155
QY      426 ThrGlySerProThrThrIleThrIleArgProPheGluSer-----TyrLysVal 442
DB      1156 TCTACTGAACCTGTACAAACCTA-----AGCTTGTATGACAAAAAGTTATGAACT 1209
QY      443 SerIleValIleAspArgGlnSerProPro-----ValSerPro 455
DB      1210 ATAGCTATATACAGAGCTAGCGCGCTTGGCGAATGTAAGTATATTAGCTTACGAA 1269
QY      456 IleGlnProHisPheIleIleAsnGlnIleGluLeuTyrLeuAsnGlySerSerAsnAsn 475
DB      1270 GTTGAT-----TTTAGTCAATATGATGATCAAAAATAAGAAACTGATACACA 1317
QY      476 ThrLeuLysTyrSerAlaGlyLysIleLeuSerAsnTyrGlnAsnThrThrPheGln 495
DB      1318 ACATATGATTCAAAAGAAACAAATGCGCATGTAATGCAAGGATTTATTTGACCAATTA 1377
QY      496 PheProArgLysAspCysAsnLeuValIleAspProGlyCysSerProAsnPheAsn 515
DB      1378 CCGCCAGAAACACAGAT-----GAACCACTTGA AAAA 1410
QY      516 AsnTyrSerHisIleLeuSerHisPheSerLeuPheThrTyrSerTyrValIleGlyLeu 535
DB      1411 GCATATAGTCATCAGCTTAATTAATTCGCGAATGTTCTTA-----ATG 1452
QY      536 GlnLeuGlnIleLeuAspThrGlyValLeuGlyTyrTrpThrHisSerSerValAspArgTyr 555
DB      1453 CAGAGCCGTGTGGAACAATTCATTTTCTTGAACACATGAAAGGTATGACATTTT 1512
QY      556 AsnAlaIleSerAspLysIleIleThrMetIleProAlaIleLysGlyAsnAsnLeuAsp 575
DB      1513 AATACAAATTTGATCGGAAAAGATTTACTCACTTCCAGTAGGAAAGCATATCCCTTGCT 1572
QY      576 ThrAsnSerLysValIleGluGlyProGlyHisThrGlyGlyAsnLeuValTyrLeuGln 595
DB      1573 TCAGGTGCTTCATATTTGAAGGTCCAGAGTTCACAGAGGAAATTTACTATCTTA 1632
QY      596 -----SerGlnGlyArgLeuGluIleThrCysGluThrProAsnSerThrGln 611
DB      1633 GAATCTAGTAATTTCAATTTGCTAAATTTAAAGTTACATTAATTAATTCAGAGCTTGTACAA 1692
QY      612 SerTyrPheIleArgLeuArgTyrAlaThrAsnGlyAlaGlyAsnThrLeuProAsn 631
DB      1693 CGATATCGGTATGAAATACGGCTATCTCTAC-----ACTAACTTA 1734
QY      632 SerLeuThrIleProGlyValIleGlyIleProProGlnArgLeuAsnAsnThrPheSer 651

```

Db 1735 CGACCTTTGTG-----CAAAATTCAAACATGATTTCTT 1770  
Qy 652 GYThrAsnTyAsn-----AsnLeuGlnTyArgLysPheGly 664  
Db 1771 GTTCATCTACATTAAATACTGATTAAGATGATTAACATTAACATTTGAT 1830  
Qy 665 TYrPheGlnPheProSerThrValThrLeuProLeuAsnArgAsnIleProPheIlePhe 684  
Db 1831 CTCGCACTACTAATTCTTAATATGGGCTCTCGGGGATAGAGATTAAGTAACTTAATAGGA 1890  
Qy 685 AsnArgAlaAspValSerAsnSerIleLeuIleIleAspLysIleGluPheIleProIle 704  
Db 1891 GCAGAACTTGTGTTCTTAATGAAAAATCTATATAGATTAAGATTTATCCAGTA 1950  
RESULT 10  
US-08-993-722A-37  
/ Sequence 37, Application US/08993722A  
/ Patent No. 6060594  
/ GENERAL INFORMATION:  
/ APPLICANT: English, Leigh H.  
/ APPLICANT: Brubock, Susan M.  
/ APPLICANT: Malvar, Thomas M.  
/ APPLICANT: Bryson, James W.  
/ APPLICANT: Kulesza, Caroline A.  
/ APPLICANT: Walters, Frederick S.  
/ APPLICANT: Staltn, Stephen L.  
/ APPLICANT: Von Teresch, Michael A.  
/ TITLE OF INVENTION: NUCLEIC ACID SEGMENTS ENCODING MODIFIED  
/ TITLE OF INVENTION: COLEOPTERAN-TOXIC CRYSTAL PROTEINS  
/ NUMBER OF SEQUENCES: 113  
/ CORRESPONDENCE ADDRESS:  
/ ADDRESSEE: Arnold, White & Durkee  
/ STREET: P.O. Box 4433  
/ CITY: Houston  
/ STATE: Texas  
/ COUNTRY: USA  
/ ZIP: 77210  
/ COMPUTER READABLE FORM:  
/ MEDIUM TYPE: Floppy disk  
/ COMPUTER: IBM PC compatible  
/ OPERATING SYSTEM: PC-DOS/MS-DOS  
/ SOFTWARE: Patentin Release #1.0, Version #1.30  
/ CURRENT APPLICATION DATA:  
/ APPLICATION NUMBER: US/08/993,722A  
/ FILING DATE: 18-DEC-1997  
/ CLASSIFICATION: 435  
/ ATTORNEY/AGENT INFORMATION:  
/ NAME: Kitchell, Barbara S.  
/ REGISTRATION NUMBER: 33,928  
/ REFERENCE/DOCKET NUMBER: MECO:149  
/ TELECOMMUNICATION INFORMATION:  
/ TELEPHONE: 512/418-3106  
/ TELEFAX: 512/474-7577  
/ INFORMATION FOR SEQ ID NO: 37:  
/ SEQUENCE CHARACTERISTICS:  
/ LENGTH: 1959 base pairs  
/ TYPE: nucleic acid  
/ STRANDEDNESS: single  
/ TOPOLOGY: linear  
/ FEATURE:  
/ NAME/KEY: CDS  
/ LOCATION: 1..1956  
/ US-08-993-722A-37  
Alignment Scores:  
Pred. No.: 2 586-60 Length: 1959  
Score: 663.00 Matches: 208  
Percent Similarity: 43.5% Conservative: 114  
Best Local Similarity: 28.1% Mismatches: 292  
Query Match: 17.1% Indels: 126  
DB: 3 Gaps: 27

US-10-783-417-2 (1-735) x US-08-993-722A-37 (1-1959)  
Qy 1 MetAsnGlnAsnAspAsnAsnGlnTyArgIleIleAspSerHisThrSerProTyR 20  
Db 1 ATGAATCCAAACAAAT---CGAAGTCAACATGAT-----ACGTAAGATT 42  
Qy 21 PheProAsnArg-----AsnSerAsnAspSerArgTyProTyThrAsnProAsn 38  
Db 43 ACACCTAACAGATGATTTGCAATCAATCAATATCTCTTACCTGACCAATCCAAAT 102  
Qy 39 GlnProLeuGlnAsnThrAsnTyArgLysGlnTyPheAsnMetCysGln---GlyAsnThr 57  
Db 103 TCACACCTAGAGAAATTAATTAATAAGATTTTAAGATGACGAGAACAGCTTCTAG 162  
Qy 58 GlnTyArgLysAsnAsnPheGlnThrPheAlaSerIleAla---AlaValSer 76  
Db 163 GAAGTCTAGACAC---TCACAGTAAAGATCCAGTTGGG 201  
Qy 77 AlaGlyThrIleValSerGlyThrLeuLeuAlaGlyIleGlyGlyLeuThrSerIleSer 96  
Db 202 ACAGAAATTTCTGTTGATAGGCGACATTTTAGGTGTAGGA----- 243  
Qy 97 GlyProIleGlyIleIleGlyAlaIleIleIleSerPheGlyThrLeuIleThrValPhe 116  
Db 244 -----GTTCCATTTGCTGGGCGACCTCATTTTATCATCATTTCTTAACACTATA 297  
Qy 117 TrpProAlaGlyGlnGlnAspLysThrValThrThrGlnPheIleLysMetGlyGluIle 136  
Db 298 TGGCCCAAGT-----GATCCGACCCATGAGAGCGCTTTATGACCAAGTTGCAAGTA 348  
Qy 137 PheValAspThrProLeuThrGlnSerIleLysGlnLeuLysLeuGlnThrLeuGlnGly 156  
Db 349 CTGATGATGATGAATAATGAGAGATGCTAAAGTAAAGCTCTTGACAGATTACGGGT 408  
Qy 157 PheArgGlnIleLeuGlnSerTyAsnThrAlaLeuAspAspTyArgLysLeuLysArg 176  
Db 409 CTTCAAAATTAATTTGGAAGATTATGTTAATGCTTAATTCCTGGAAGAA----- 459  
Qy 177 LeuGlnAlaProGlyLeuProProSerSerAlaLeuGlnAlaAlaLeuThrLeuLys 196  
Db 460 -----ACACCTTAAGTTGCGAATGAAGAAAGCAAGATGCAATGAAG 504  
Qy 197 IleArgPheGlnAsnValHisAsnAspPheIleArgGluIleProGlyPheGlnLeuGln 216  
Db 505 GAACCTTTTCTCAAGCAGAAAGTCATTTGTAATTCATCCATCCGTCATTTCCAGCTTCC 564  
Qy 217 ThrTyArgThrLeuLeuLeuProIleTyArgAlaGlnAlaAlaAsnPheHisLeuAsnLeu 236  
Db 565 AAATTCGAAGTCTGTTTCTACCAACATATGACACAGCTGCAAAATACATTTATTTGCTA 624  
Qy 237 LeuGlnGlnGlyAlaGlnLeuAlaAspGlnTyPheAsnAlaAspIleHisProSerGlnIle 256  
Db 625 TTAAGAAAGATGCTCAAGTTTTCGAGAGAAATGGGATATTC----- 666  
Qy 257 GlnProAlaAlaGlyThrSerAspAspTyArgLysLeuLeuLysGlnAsnIleProLys 276  
Db 667 -----TCAGAAAGATGTTGCTGAATTTATCATACATTAACAATTAACTTACA---CAACAA 717  
Qy 277 TySerAsnTyArgAlaAsnThrTyArgThrGlyLeuLysAsnLeuArgAspGlnPro 296  
Db 718 TACACTGACCATTTGCTTAATTTGATTAATGTTGAATTAATGCTTTAAGAGTTCAACT 777  
Qy 297 AsnMetLysTrpSerIlePheAsnAspTyArgArgTyMetThrIleThrValLeuAsp 316  
Db 778 TATGATGATGGGTCAAAATTTAACCCGTTTTCGACAGAAATGACCTTAACCTGATTAAGAT 837  
Qy 317 ThrIleSerGlnPheSerLeuTyArgAspIleLysArgTyArgAspSerIleGlyGlyIle 336  
Db 838 CTAAATTTGACTTTTCCATTTATATATATCGGTTATCTCA----- 879  
Qy 337 GluValLysGlyIleLysAsnGlnLeuThrArgGluIleTyThrThrGlnIleAsnPhe 356

Db 880 -----AAGGGGTTAAAAAGAGACTAACAGACATTTTTCAGATCCAAATTTTCT 933  
Qy 357 AsparGleuProGlnLeuArgValGlnProAsnLeuAlaThrMetGluTyrAsnLeuThr 376  
Db 924 CTTAGACACCA---CTTGCGTAGAGCACTTTTTCAGATGAGTAAGAAACCTCATTTGCA 990  
Qy 377 ArgAlaSerPheLeuPheSerPheLeuGlnGlnPheLeuPheTyrThr----- 393  
Db 991 AAACCT-----CATTTATGATATATTTACAGGGAGATGGAATTTTCATACGGCTCTGCA 1044  
Qy 394 -----GluAsnThrAsnHe-----GluAsnArgLeuValGlyLe 405  
Db 1045 CTGGTTACTTTGGAAAGATTTCTTCATATTTGCTGCTGTAATAT-----GTA 1095  
Qy 406 SerAsnArgAspAlaProThrTyrSerAsnThrIleThrGlnThrLeuTyrGlyGlnArg 425  
Db 1096 GAAACTAGACCTAGATAGATGATCTAGTAAGACATTACTTCCCATTTTATGAGATATAA 1155  
Qy 426 ThrGlySerProThrThrIleThrIleArgProPheGluSer-----TyrIleVal 442  
Db 1156 TCTACTGAACCTGTACAAAAGCTA-----ACCTTGTAGTACAAAAGTTTATCGAACT 1209  
Qy 443 SerIleValThrAspArgGlnSerProPro-----ValSerPro 455  
Db 1210 ATAGCTAATACAGAGTAGCGGCTTGCGCAATGTAAGTATTTAGTGTACGAAA 1269  
Qy 456 IleGlnProHisPheIleIleAsnGlnIleGlnLeuTyrLeuAsnGlySerSerAsnAsn 475  
Db 1270 GTTGAT-----TTTAGTCATATGATGATGATCAAAAAGAAAGAACTAGTACACA 1317  
Qy 476 ThrLeuLysTyrSerAlaGlyGlySerLeuSerAsnTyrGlnAsnThrPhePheGln 495  
Db 1318 ACATGTGATTCAAAAGAAACAATGCGCATGTAGTACAGAGATTTTATGACCAATTA 1377  
Qy 496 PheProArgGlyLysAspCysAsnLeuValIleAspProGlyCysSerProAsnHeAsn 515  
Db 1378 CCGCGAAGAAACAAGAT-----GAACCACTTGAAAAA 1410  
Qy 516 AsnTyrSerHisIleLeuSerHisPheSerLeuPheThrTyrSerTyrValIleGlyLeu 535  
Db 1411 GCATATAGCATCAGCTTAATTAACCGCAATGTTCTTA-----ATG 1452  
Qy 536 GlnLeuGlnIleLeuAspThrGlyValLeuGlyTyrPheHisSerSerValAspArgGly 555  
Db 1453 CAGGACCGCTGGAACAATTCATTTTACTTGACACATAGTAAGTGTAGACTTTT 1512  
Qy 556 AsnAlaIleSerAspLysIleIleThrMetIleProAlaIleIleGlyAsnAsnLeuAsp 575  
Db 1513 AATACAAATTTGATGCTGAAGAAATTACTCAACTTCAGTAGTGAAGCATATGCTTGTCT 1572  
Qy 576 ThrAsnSerLysValIleGlnGlyProGlyHisThrGlyGlyAsnLeuValTyrLeuGln 595  
Db 1573 TCAAGTGTCTTCATTTTGAAGTCCAGATTCACAGAGAAATTTACTATCTCTAAAA 1632  
Qy 596 -----SerGlnGlyArgLeuGlnIleThrCysGlnThrProAsnSerThrGln 611  
Db 1633 GAATCTAGTAATTCATTCCTAAATTAATTAAGTTAAATTCAGACAGCTTGTTACAA 1692  
Qy 612 SerTyrPheIleArgLeuArgTyrAlaThrAsnGlyAlaGlyAsnThrLeuProAsnIle 631  
Db 1693 CGATATCGTGAAGATAGCTATGCTTCTTACC-----ACTAATCTTA 1734  
Qy 632 SerLeuThrIleProGlyValIleGlyIleProProGlnArgLeuAsnAsnThrPheSer 651  
Db 1735 CGACTTTTGTG-----CAAAATTCAAACAATATTTTCT 1770  
Qy 652 GlyThrAsnTyrAsn-----AsnLeuGlnTyrGlyAspPheGly 664  
Db 1771 GTCATCTACATTAATAAATCTAGATAAAGATGATTAACATATCAACATTTGAT 1830  
Qy 665 TyrPheGlnPheProSerThrValThrLeuProLeuAsnArgAsnIleProHeIlePhe 684  
Db 1831 CTGCAACTACTAATTTATATATGAGGTTCTCGGGGTAGTAAGATGAACCTTATATAGGA 1890

Qy 685 AsnArgAlaAspValSerAsnSerIleLeuIleIleAspLysIleGlnPheIleProIle 704  
Db 1891 GCAGAACTCTGCTTCTTAATGAAGAAATCTATATAGATGAAATTTATCCAGATA 1950  
RESULT 11  
US-08-993-170A-37  
Sequence 37, Application US/08993170A  
Patent No. 6063597  
GENERAL INFORMATION:  
APPLICANT: English, Leigh H.  
APPLICANT: Brunsack, Susan M.  
APPLICANT: Malvar, Thomas M.  
APPLICANT: Bryson, James W.  
APPLICANT: Kulesza, Caroline A.  
APPLICANT: Walters, Frederick S.  
APPLICANT: Slater, Stephen L.  
APPLICANT: Von Terach, Michael A.  
TITLE OF INVENTION: POLYPEPTIDE COMPOSITIONS TOXIC TO  
NUMBER OF SEQUENCES: 113  
CORRESPONDENCE ADDRESSES:  
ADDRESSER: Arnold, White & Durkee  
STREET: P.O. Box 4433  
CITY: Houston  
STATE: Texas  
COUNTRY: USA  
ZIP: 77210  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/993,170A  
FILING DATE: 18-DEC-1997  
CLASSIFICATION: 424  
ATTORNEY/AGENT INFORMATION:  
NAME: Kitchell, Barbara S.  
REGISTRATION NUMBER: 33,928  
REFERENCE/DOCKET NUMBER: MECO:002  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 512/474-7577  
TELEFAX: 512/418-3000  
INFORMATION FOR SEQ ID NO: 37:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1959 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 1..1956  
US-08-993-170A-37  
Alignment Scores:  
Pred. No.: 2.58e-60 Length: 1959  
Score: 663.00 Matches: 208  
Percent Similarity: 43.5% Conservative: 114  
Best Local Similarity: 28.1% Mismatches: 292  
Query Match: 17.1% Indels: 126  
DB: 3 Gaps: 27  
US-10-783-417-2 (1-735) x US-08-993-170A-37 (1-1959)  
Qy 1 MetAsnGlnAsnAsnAspAsnAsnGlnTyrGlnIleIleAspSerHisThrSerProTyr 20  
Db 1 ATGAATCCAAACAAT---CGAAGTGACATGAT-----ACGATTAAGGTT 42  
Qy 21 PheProAsnArg-----AsnSerAsnAspSerArgTyrProTyrThrAsnAsnProAsn 38  
Db 43 ACACCTTAACAGTGAATTCGAAACGTAACATATATCTTTTACGTGACATCCAAAT 102

OY	39	GLNProLeuGGAAaenThrAsnTYrLeuGluTYrLeuAsnMetCysGln---	GLYAsnThr	57
Db	103	TCAACACATGAAAGAAATTAATTAATAAGAAATTTTAAAGAAATGACTGAAGACAGTTCTACG		162
OY	58	GLNTrpGLYAspAsnPhgGluThrPheAlaSerAlaMetThrIleAla---	AlaValSer	76
Db	163	GAGGTGCTAGAACAC-----TCACAGTAAAGATGCACTTGG		201
OY	77	AlaGlyThrIleValSerGlyThrLeuAlaGlyIleGlyGlyLeuThrSerIleSer		96
Db	202	ACAGGAATTTCTGTTGTAGAGGACAGATTTTACGTGTGTAGGA-----		243
OY	97	GlyProIleGlyIleIleGlyAlaIleIleIleSerPheGlyThrLeuIleThrValPhe		116
Db	244	-----GTTCCATTGTCGGGACCTCTCATTTTATTCATCATTTCTTAAACCTATA		297
OY	117	TrpProIaGlyGluGlnAspLysThrValTrpThrGlnPheIleLysMetGlyIle		136
Db	288	TGGCCAACT-----GATGCTACCCATGGAAGGCTTTATGGCACAACTGAAGTA		348
OY	137	PheValAspThrProLeuThrGluSerIleLysGlnLeuLysLeuGlnThrLeuGluGly		156
Db	349	CTGATAGATTAAGAAATATGAGAGATGATCTAAAGTAAAGCTCTTGACAGATTAACAGGT		408
OY	157	PheArgGlnIleLeuGlnSerTYrAsnThrAlaLeuAspAspTrpArgLysLeuLysArg		176
Db	409	CTTCAAAATTAATTTGAAAGATTTATTTAATGGGTAAATCTCTGGAAGAAA-----		459
OY	177	LeuGlnAlaProGlyLeuProProSerSerAlaLeuGlnGlnAlaIleThrLeuLys		196
Db	460	-----ACACCTTAAAGTTTGGCAAGTAAAGAACCCAAAGATGCAATTAAGG		504
OY	197	IleArgPheGluAsnValHisAsnAspPheIleArgGluIleProGlyPheGlnLeuGlu		216
Db	505	GAACTTTTTCACAGCAAGAAAGTCATTTTGATATTCATGCTCCGTCATTTGCAGTTTCC		564
OY	217	ThrTYrLysThrLeuLeuLeuProIleTYrAlaGlnAlaAlaAsnPheHisLeuAsnLeu		236
Db	565	AAATTCGAAGTCTGTCTTCTACCAACATATGCAACAGCTGCMAATTAACATTATTTGCTA		624
OY	237	LeuGlnGlnGlyAlaGluLeuAlaAspGluTrpAsnAlaAspIleHisProSerGlnIle		256
Db	625	TTAAAGATGCTCAAGCTTTTGGAGAAAGATGGGATTTCT-----		666
OY	257	GluProAsnAlaGlyThrSerAspAspTYrTYrLysLeuLeuLysGluAsnIleProLys		276
Db	667	-----TCGAAGAAGTGTCTGAATTTATTCATACACATTTAAACTTACA---CAACAA		717
OY	277	TYrSerAsnTYrCysAlaAsnThrTYrArgThrGlyLeuLysAsnLeuArgAspGluPro		296
Db	718	TACACTGACCATGTGGTTAATGTATATATGTGGATTAAATGGTTTAAAGAGTTCAACT		777
OY	297	AsnMetLysTrpSerIlePheAsnAspTYrArgArgTYrMetThrIleThrValLeuAsp		316
Db	778	TATGATGATGGGTCAAAATTAAACGGTTTGGCAAGAAATGACTTTAACTGATTAATGAT		837
OY	317	ThrIleSerGlnPheSerLeuTYrAspIleLysArgTYrArgAspSerIleGlyIle		336
Db	838	CTAATTTGACTTTTCCATTTATATATATCGGTATACTCA-----		879
OY	337	GluValLysGlyIleLysAsnGluLeuThrArgGluIleTYrThrThrGluIleAsnPhe		356
Db	880	-----AAAGGGGTTTAAACAGAACTTAAACAGACAGCATTTTAAACGATTCATTTTCT		933
OY	357	AspArgLeuProGlnLeuArgValGlnProAsnLeuAlaThrMetGluTYrAsnLeuThr		376
Db	934	CTTAGACACCA---CTTGCGTAGCGAACCACTTTTGGAGTATAGAAACCTATTCGA		990
OY	377	ArgAlaSerPheLysLeuPheSerPheLeuGlnGlnPheIlePheTYrThr-----		393
Db	991	AAACT-----CATTATTTGATTAATTATTCACGGGAGTTGATTTTCATACGGGTCTTCA		1044
OY	394	-----GluAsnThrAsnPhe-----GlyAsnArgLeuValGlyIle		405

[illegible]

RESULT 12  
US-08-993-775B-37  
; Sequence 37, Application US/08993775EE  
; Patent No. 6077824  
; GENERAL INFORMATION:  
; APPLICANT: English, Leigh H.

APPLICANT: Brusock, Susan M.  
APPLICANT: Malvar, Thomas M.  
APPLICANT: Bryson, James W.  
APPLICANT: Kulesza, Caroline A.  
APPLICANT: Walters, Frederick S.  
APPLICANT: Slatin, Stephen L.  
APPLICANT: Von Terach, Michael A.  
TITLE OF INVENTION: METHODS FOR IMPROVING THE ACTIVITY OF  
NUMBER OF SEQUENCES: 113  
TITLE OF INVENTION: DELTA-ENDOTOXINS AGAINST INSECT PESTS  
CORRESPONDENCE ADDRESS:  
ADDRESSES: Arnold, White & Durkee  
STREET: P.O. Box 4433  
CITY: Houston  
STATE: Texas  
COUNTRY: USA  
ZIP: 77210  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/993,775B  
FILING DATE: 18-DEC-1997  
CLASSIFICATION: 514  
ATTORNEY/AGENT INFORMATION:  
NAME: Kitchell, Barbara S.  
REGISTRATION NUMBER: 33,928  
REFERENCE/DOCKET NUMBER: MCO:150  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 512/418-3000  
TELEFAX: 512/474-7577  
INFORMATION FOR SEQ ID NO: 37:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1959 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 1..1956  
US-08-993-775B-37  
US-10-783-417-2 (1-735) x US-08-993-775B-37 (1-1959)  
Alignment Scores:  
Pred. No.: 2,586-60 Length: 1959  
Score: 663.00 Matches: 208  
Percent Similarity: 43.5% Conservative: 114  
Best Local Similarity: 28.1% Mismatches: 292  
Query Match: 17.1% Indels: 126  
DB: 3 Gaps: 27

QY 97 GlyProIleGlyIleIleGlyAlaIleIleIleSerPheGlyThrLeuIleThrValPhe 116  
DB 244 -----GTTCCATTGCTGGGGGCACTCACTTCATTATTCATCATCTTCTTAACACTATA 297  
QY 117 ThrProAlaGlyGluGlnAspLeuThrValIleThrGlnPheIleIleYemecGlyGlyIle 136  
DB 298 TGGCCAACT-----GATGCTGACCCCAAGAAAGGCTTTTGGCAACAAGTTAAAGTA 348  
QY 137 PheValAspThrProLeuThrGluSerIleIleGlyLeuIleLeuGlnThrLeuGly 156  
DB 349 CTGATAGTAAAGAAATAGAGAGATGCTAAAGTAAGCTCTTGCAAGATTACAGGGT 408  
QY 157 PheArgGlnIleLeuGlnSerTyraThrAlaLeuAspAspTrpArgLeuLeuArg 176  
DB 409 CTTCAAAATTAATTCGAAGATTATGTTAATCGTTAAATTCCTGGAAGAAA----- 459  
QY 177 LeuGlnAlaProGlyLeuProProSerSerAlaLeuGlnGlnAlaIleLeuThrLeu 196  
DB 460 -----ACACCTTAAGTTTGGCAAGTAAGAAAGCAAGATCGAATTAAG 504  
QY 197 IleArgPheGluAsnValIleAsnAspPheIleArgGluIleProGlyPheGlnLeuGlu 216  
DB 505 GAACCTTTTCTCAAGCAGAAAGTCATTTTGTATTCATGCCATTCGAGTTCC 564  
QY 217 ThrTyrlsThrLeuLeuLeuProIleTyraGlnAlaAlaAsnPheIleLeuAsnLeu 236  
DB 565 AAATCGAAGTGTGTTCTACCAACATATGCACAAAGTCGCAAAATACATTTATGCTA 624  
QY 237 LeuGlnGlnGlyAlaGluLeuAlaAspGlyTrpAsnAlaAspIleIleProSerGlnIle 256  
DB 625 TTAAGAGTGTCAAGTTTGGAGAAAGAAAGGAGATTTCT----- 666  
QY 257 GluProAsnAlaGlyThrSerAspAspTyrlsTyrlsLeuLeuGluAsnIleProIys 276  
DB 667 -----TCAGGAAGATGTTGCGAATTTTATCATAGACATTTAAACTTACA---CAACAA 717  
QY 277 TyrlsAsnTyrlsValAsnThrTyraGlyLeuLeuIleLeuIleLeuAspGluPro 296  
DB 718 TACACTGACCATGTTGTTAATGTTAATGTTAATGTTAATGTTAAGAGTTCAACT 777  
QY 297 AsnMetIysTrpSerIlePheAsnAspTyrlsArgGlyTrpMetThrIleThrValLeuAsp 316  
DB 778 TATGATGCATGCTCAAAATTTAACCGTTTGGCAAGAAATGACTTATGATATTAAT 837  
QY 317 ThrIleSerGlnPheSerLeuTyraPheIleIleIleIleIleIleIleIleIleIle 336  
DB 838 CTAATTTGACTTTTCCATTATATATTCGTTATCTCA----- 879  
QY 337 GluValIleGlyIleIleAsnGluLeuThrArgGluIleTyrlsThrGluIleAsnPhe 356  
DB 880 -----AAAGGGGTTAAACAGAACTAACAGAGACATTTTACGATCAATTTTCT 933  
QY 357 AspArgLeuProGlnLeuArgValGlnProAsnLeuAlaThrMetGluTyraAsnLeuThr 376  
DB 934 CTTAGGACACCA---CTTGCGTACCGAACACTTTTGTAGATATGAAAACTATTCGA 990  
QY 377 ArgAlaSerPheIlePheSerPheLeuGlnPheIlePheTyrlsThr----- 393  
DB 991 AAACCT-----CATTTATGATTATTTACAGGGGATTTGAATTTTCATACCGCTTCGA 1044  
QY 394 -----GluAsnThrAsnPhe-----GlyAsnArgLeuValGlyIle 405  
DB 1045 COTGTTACTTTGGGAAGATTTCTTCAATATATGCTCGGTAATAT-----GTA 1095  
QY 406 SerAsnArgAspAlaProThrTyrlsAsnThrIleThrGluThrLeuTyrlsGlyArg 425  
DB 1096 GAAACTAGACCTAGATAGATCTAGTAAGCAATATCTTCCCATTTTATAGAGATAAA 1155  
QY 426 ThrGlySerProThrThrIleThrThrIleArgProPheGlySer-----TyrlsVal 442  
DB 1156 TCTACTGAACCTGTAACAAAGCTA-----AGCTTGTATGACAAAAAGTTTATGAACT 1209





Db 298 TGGCCAGT-----GAGCTGACCCAGTGAAGGCTTTATGAGCAACAGTTGAAGTA 348  
Qy 137 PheValAspThrProLeuThrGluSerIleValGlnLeuLeuLeuGlnThrLeuGlnGly 156  
Db 349 CTGATAGATTAAGAAATAGAGATATGCTAAAGTAAAGCTCTTGCAAGTATACAGGT 408  
Qy 157 PheArgGlnIleLeuGlnSerTyrAsnThrAlaLeuAspAspTyrArgIleValArg 176  
Db 409 CTTCAAAATTAATTTGCAAGATTATGTTAATGCTTAAATTCCTGGAGAGAAA----- 459  
Qy 177 LeuGlnAlaProGlyLeuProProSerSerAlaLeuGlnGlnAlaAlaLeuThrLeuLys 196  
Db 460 -----ACACCTTTAAGTTTGGCAAGTAAAGAACCAAGATCGAAATAGG 504  
Qy 197 IleArgPheGluAsnValHisAsnAspPheIleArgGlnIleProGlyPheGlnLeuGln 216  
Db 505 GAACCTTTTCTCAAGCAGAAAGTCTATTGCTAATTCATGCCGTCATTTGCAAGTTCC 564  
Qy 217 ThrTyrLeuThrLeuLeuLeuProIleTyrAlaGlnAlaAlaAsnPheHisLeuAsnLeu 236  
Db 565 AAATTCGAAGTCTGTTCTTACCAACATATGCAACAAGCTGCAAAATATACATTTATGCTA 624  
Qy 237 LeuGlnGlnGlyAlaGlnLeuAlaAspGluTyrAsnAlaAspIleHisProSerGlnIle 256  
Db 625 TTTAAAGATGCTCAAGTTTGGAGAGAGATGGGATATCT----- 666  
Qy 257 GluProAsnAlaGlyThrSerAspAspTyrTyrTyrLeuLeuLeuGlnAsnIleProLys 276  
Db 667 -----TCGAAGATGTTGCTGAATTTTATCATACATTAATACTTACA---CAACA 717  
Qy 277 TyrSerAsnTyrCysAlaAsnThrTyrArgThrGlyLeuLeuAsnLeuArgAspGluPro 296  
Db 718 TACACGTGACCATGTTGTTAATGTAATGTTGATTAATGTTAATGTTAAGAGTTCAACT 777  
Qy 297 AsnMetLysTyrSerIlePheAsnAspTyrArgArgTyrMetThrIleThrValLeuAsp 316  
Db 778 TATGATGATGGGTCAAAATTTAACCGTTTTCGACAGAAATATGACTTATATTAAT 837  
Qy 317 ThrIleSerGlnPheSerLeuTyrAspIleLysArgTyrArgAspSerIleGlyGlyIle 336  
Db 838 CTAAATTTGATCTTTCCCATTTATATATATTCGGTATACTCA----- 879  
Qy 337 GluValIleGlyIleLysAsnGlnLeuThrArgGlnIleTyrThrThrGlnIleAsnPhe 356  
Db 880 -----AAAGGGGTAAACAGAACATCAACAGACATTTTACGATCCAAATTTTCT 933  
Qy 357 AspArgLeuProGlnLeuArgValGlnProAsnLeuAlaThrMetGluTyrAsnLeuThr 376  
Db 934 CTTAGGACACCA---CTTGGTACGACCACTTTTGTGATATGAAAACTCTATTCGA 990  
Qy 377 ArgAlaSerPheLysLeuPheSerPheLeuGlnGlnPheIlePheTyrThr----- 393  
Db 991 AAACCT-----CATTTATTTGATATTTTACAGGGATTTGAATTTTCATACGGCTTCA 1044  
Qy 394 -----GluAsnThrAsnPhe-----GlyAsnArgLeuValGlyIle 405  
Db 1045 CTGTGTTACTTTGGAAAGATTTCTTCAATATATGCTGTAATAT-----GTA 1095  
Qy 406 SerAsnArgAspAlaProThrTyrSerAsnThrIleThrGlnThrLeuTyrGlyGluArg 425  
Db 1096 GAACTTACGCTTATAGATCTAGTAAAGACATTACTCCCATTTTATAGAGATTA 1155  
Qy 426 ThrGlySerProThrThrIleThrIleArgProPheGlnSer-----TyrLysVal 442  
Db 1156 TCTACTGAACCTGTACAAAAGCTA-----ACCTTTGATGACAAAAAGTTATGCAACT 1209  
Qy 443 SerIleValThrAspArgLysSerPro-----ValSerPro 455  
Db 1210 ATAGCTAATACAGAGCTAGCGCTGGCGGAATGGTAAGTATATTAGCTTTACGAAA 1269  
Qy 456 IleGlnProHisPheIleIleAsnGlnIleGluLeuTyrLeuAsnGlnSerSerAsnAsn 475  
Db 1270 GTTGAT-----TTTAGTCATATGATGATCAAAAAAATGAACTAGTACACA 1317

Qy 476 ThrLeuLysTyrSerAlaGlyGlySerLeuSerAsnTyrGlnAsnThrThrPhePheGln 495  
Db 1318 ACATATGATTTCAAAAAGAAACAATGCCATGTAAGTACAGACAGATTTCTATTCACCAATTA 1377  
Qy 496 PheProArgLysLysAspCysAsnLeuValIleAspProGlyCysSerProAsnPheAsn 515  
Db 1378 CCGCCAGAAACAACAGAT-----GAACCACTTGAAAAA 1410  
Qy 516 AsnTyrSerHisIleLeuSerHisPheSerLeuPheThrTyrSerTyrValIleGlyLeu 535  
Db 1411 GCATATATGATCATCAGCTTAATTAATTCGCGGAATGTTTCTTA-----ATG 1452  
Qy 536 GlnLeuGlnIleLeuAspThrGlyValLeuGlyTyrThrHisSerSerValAspArgTyr 555  
Db 1453 CAGAGCCGTCGTGGAACAATTCATTTTACTTGACACATAGAAAGTATACATTTT 1512  
Qy 556 AsnAlaIleSerAspLysIleIleThrMetIleProAlaIleLysGlyAsnAsnLeuAsp 575  
Db 1513 AATACAAATTTGATGCTGAAAGATTACTCACTTCCAGTATGTAAGCATATGCCCTGTCT 1572  
Qy 576 ThrAsnSerLysValIleGlnGlyProGlyHisThrGlyLysAsnLeuValTyrLeuGln 595  
Db 1573 TCAAGTCTTCATTTATGAAAGTCCAGATTCACAGAGAAATTTACTATTCCTTAAA 1632  
Qy 596 -----SerGlnGlyArgLeuGlnIleThrCysGluThrProAsnSerThrGln 611  
Db 1633 GAATCTAGTAATTCATATGCTAAATTAATTAAGTATCAATTAATTCAGCAGCCTTGTTACA 1692  
Qy 612 SerTyrPheIleArgLeuArgTyrAlaThrAsnGlyAlaGlyAsnThrLeuProAsnIle 631  
Db 1693 CGATATCGTGAAGATAGCTATGCTCTTCC-----ACTAATCTTA 1734  
Qy 632 SerLeuThrIleProGlyValIleGlyIleProProGlnArgLeuAsnAsnThrPheSer 651  
Db 1735 GCACCTTTTGG-----CAAAATTCAAACATATTTTCTT 1770  
Qy 652 GlyThrAsnTyrAsn-----AsnLeuGlnTyrGlyAspPheGly 664  
Db 1771 GTCATCTACATTAATTAAGATATGATGAATAAGATGATGATTTTAACTATCAAACTTTGAT 1830  
Qy 665 TyrPheGlnPheProSerThrValThrLeuProLeuAsnAspAsnIleProPheIlePhe 684  
Db 1831 CTGCACTACTAATTTCAATATATGAGGTCTCGGGTGAATAAGAACTTATTAATAGGA 1890  
Qy 685 AsnArgAlaAspValSerAsnSerIleLeuIleIleAspLysIleGluPheIleProIle 704  
Db 1891 GCAGATCTCTTCGTTCTTCAATGAAAAAATCTATATAGATAGATTAATTCACAGTA 1950

RESULT 14  
US-09-427-769-37  
Sequence 37, Application US/09427769  
Patent No. 6642030  
GENERAL INFORMATION:  
APPLICANT: English, Leigh H.  
APPLICANT: Brusock, Susan M.  
APPLICANT: Malvar, Thomas M.  
APPLICANT: Bryson, James W.  
APPLICANT: Kuleza, Caroline A.  
APPLICANT: Walters, Frederick S.  
APPLICANT: Slatin, Stephen L.  
APPLICANT: Von Tersch, Michael A.  
APPLICANT: Romano, Charles  
TITLE OF INVENTION: NUCLEIC ACID SEGMENTS ENCODING MODIFIED  
TITLE OF INVENTION: COLEOPTERAN-TOXIC CRYSTAL PROTEINS  
NUMBER OF SEQUENCES: 113  
CORRESPONDENCE ADDRESS:  
ADDRESSER: Arnold, White & Durkee  
STREET: P. O. Box 4433  
CITY: Houston  
STATE: Texas  
COUNTRY: USA  
ZIP: 77210

```

COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/427,769
FILING DATE:
CLASSIFICATION:
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: 08/993,722
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Kitchell, Barbara S.
REGISTRATION NUMBER: 33,928
REFERENCE/DOCKET NUMBER: MECO:149
TELECOMMUNICATION INFORMATION:
TELEPHONE: 512/418-3106
TELEFAX: 512/474-7577
INFORMATION FOR SEQ ID NO: 37:
SEQUENCE CHARACTERISTICS:
LENGTH: 1959 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
FEATURE:
NAME/KEY: CDS
LOCATION: 1..1956
US-09-427-769-37

Alignment Scores:
Pred. No.: 2.58e-60 Length: 1959
Score: 663.00 Matches: 208
Percent Similarity: 43.5% Conservative: 114
Best Local Similarity: 28.1% Mismatches: 292
Query Match: 17.1% Indels: 126
Gaps: 27
DB: 3

US-10-783-417-2 (1-735) x US-09-427-769-37 (1-1956)
QY 1 MetAAGlnAsnAspAsnAsnGluTyrGluIleIleAspSerHisThrSerProTyr 20
DB 1 ATGATTCACAAACAT---CGAAGTGAACATGAT-----ACGATTAAGCTT 42
QY 21 PheProAsnArg-----AsnSerAsnAspSerArgTyrProTyrThrAsnAsnProAsn 38
DB 43 ACACCTTAACAGTGAATTCGAACCTAATCAATATCTTACCTGACCAATCCAAAT 102
QY 39 GlnProLeuGlnAsnThrAsnTyrIlysgIuTrrPleuAsnMetCysGln---GlyAsnThr 57
DB 103 TCACACCTAGGAAGAAATTAAATTAAGAATTTTAAAGAATGACCTGAAGACAGTTCTAG 162
QY 58 GlnTyrGlyAspAsnAspPheGluThrPheAlaSerAlaAspThrIleAla---AlaValSer 76
DB 163 GAAGTCCTAGACAC-----TCTACAGTAAAGATGACAGTTGGG 201
QY 77 AlaGlyThrIleValSerGlyThrIleuLeuAlaGlyIleGlyGlyLeuThrSerIleSer 96
DB 202 ACAGGAATTTCTGTTAGGGGAGATTTTAGGCTTAGGA----- 243
QY 97 GlyProIleGlyIleIleGlyAlaIleIleIleSerPheGlyThrIleuIleThrValPhe 116
DB 244 -----GTTCCATTGCTGGGCGACCTTCATTTTATCAATTCATCTTTAACTATA 297
QY 117 TrpProAlaGlyGluGlnAspLysThrValTrrThrGlnPheIleLysMetGlyGluIle 136
DB 298 TGGCCACAGT-----GATGTCGACCCATGGAAGGCTTTATGCGACAAAGTGAAGTA 348
QY 137 PheValAspThrProLeuThrGlnSerIleLysGlnLeuLysLeuGlnThrLeuGlnGly 156
DB 349 CTGATGATTAAGAAATTAAGAGATGCTAAAGAAATGAAGCTCTTCAAGATTAACAGG 408
QY 157 PheArgGlnIleLeuGlnInsertYrAsnThrAlaLeuAspAspTrpArgLysLeuValG 176

```

```

DB 409 CTTCAAAATAATTTCGAAGATTAATGTTAATGCTTAATTCCTCGAAGAAA----- 459
QY 177 LeuGlnAlaProGlyLeuProProSerSerAlaLeuGlnGlnAlaAlaLeuThrLeuLys 196
DB 460 -----ACACCTTTAAGTTTGGCAAGTTCGAAATTAAGAAAGACAGATCGAATTAAG 504
QY 197 IleArgPheGluAsnValHisAsnAspPheIleArgGluIleProGlyPheGlnLeuGln 216
DB 505 GAACCTTTTTCGACAGCAGAAAGATCTTTTGTGTAATTCATGCGCTGATTTGCACTTCC 564
QY 217 ThrTyrLysThrLeuLeuLeuProIleTyrAlaGlnAlaAlaAsnPheHisLeuAsnLeu 236
DB 555 AAATTCGAAGTCTGTTCTTACCAACATATGCAAGCTGCAAAATACATTTATATGCTA 624
QY 237 LeuGlnGlnGlyAlaGluLeuAlaAspGluTrrPaaAlaAspIleHisProSerGlnIle 256
DB 625 TTTAAAGATGCTCAAGTTTTCGAGAGAAATGGGATATCTC----- 666
QY 257 GluProAlaAlaGlyThrSerAspAspTyrTyrLysLeuLeuLysGluAsnIleProLys 276
DB 667 -----TCAGAAAGATGTTGCTGAATTTTATCATGACATTAATACTTACA---CAACA 717
QY 277 TyrSerAsnTyrCysAlaAsnThrTyrArgThrGlyLeuLysAsnLeuArgAspGluPro 296
DB 718 TACACTGACACATGTTGTTAATGATATATGTTAATGTTAATGTTAAGAGTTCAACT 777
QY 297 AsnMetLysTrpSerIlePheAsnAspTyrArgArgTyrMetThrIleThrValLeuAsp 316
DB 778 TATGATGATGGGTCAAAATTTAAACCGTTTCCGACAGAAATGACCTTAACCTGATTAAGAT 837
QY 317 ThrIleSerGlnPheSerLeuLysAspIleLysArgTyrArgAspSerIleGlyIle 336
DB 838 CTAATTTGATCTTTCCTATTATATATATTCGCTTATACTCA----- 879
QY 337 GluValLysGlyIleLysAsnGluLeuThrArgGluIleTyrThrThrGluIleAsnPhe 356
DB 880 -----AAAGCGGTAAACAGACTAACAGACATTTTACGATCCATTTTTCCT 933
QY 357 AspArgLeuProGlnLeuArgValGlnProAsnLeuAlaThrMetGluTyrAsnLeuThr 376
DB 934 CTTAGGACACCA---CTTGCGTACGACCAACTTTTTCAGATGATGAATAACTCATTTCCA 990
QY 377 ArgAlaSerPheLysLeuPheSerPheLeuGlnGlnPheIlePheTyrThr----- 393
DB 991 AAACCT-----CATTTATTTGATTAATTTTACAGGGGATTTGAATTCATACGGCTTCA 1044
QY 394 -----GluAsnThrAsnPhe-----GlyAsnArgLeuValGlyIle 405
DB 1045 CTTGCTTACTTTGGGAAAGATTTCTTCAATTTATTTGCTCTGTAATAT-----GTA 1095
QY 406 SerAsnArgAspAlaProThrTyrSerAsnThrIleThrGlnThrLeuTyrGlyGluArg 425
DB 1096 GAACCTAGACCTAGATAGATCTAGTAAGACAATTAATCTCCCAATTTATGAGATTA 1155
QY 426 ThrGlySerProThrThrLysThrIleArgProPheGlnSer-----TyrLysVal 442
DB 1156 TCTACTGAACTGTGCAAAAGCTA-----ACCTTGAAGACAAAAAGTTTATCAACT 1209
QY 443 SerIleValThrAspArgGlnSerPro-----ValSerPro 455
DB 1210 ATAGCTAATACAGACGTAGCGGCTTGCCGAATGTAAGTATATTTAGGTGTACGAA 1269
QY 456 IleGlnProHisPheIleIleAsnGlnIleGluLeuTyrLeuAsnGlySerSerAsnAsn 475
DB 1270 GTTGAT-----TTTAGTCATATGATGATCAAAAAAAGAACTAGTACACAA 1317
QY 476 ThrLeuLysTyrSerAlaGlyLysSerLeuSerAsnTyrGlnAsnThrThrPheGln 495
DB 1318 ACATATGATTCAAAAGAAACAATGCCATGTAAAGTGCACAGATTTCTATTGACCAATTA 1377
QY 496 PheProArgLysLysAspCysAsnLeuValIleAspProGlyCysSerProAsnPheAsn 515

```

1378 CCGCCAGAAACACAGAT-----GAACTGTGAAAA 1410  
QY AsnTyrSerHisIleLeuSerHisPheSerLeuPheThrTyrSerTyrValIleGlyLeu 535  
DB 1411 GCATATAGTACACCTTAATTAACGGGAATTTCTTA-----ATG 1452  
QY 536 GlnLeuGlnIleLeuAspThrGlyValLeuGlyTyrPheHisSerSerValAspArgTyr 555  
DB 1453 CAGGACCGCTCGTGAACAAATTCATTTTACTGTGACACATAGAGGTAGACTTTT 1512  
QY AsnAlaIleSerAspLysIleIleThrMetIlePheAlaIleGlyValAsnAsnLeuAsp 575  
DB 1513 AATACAAATTAATGCTGAAAAGATTACTCACTTCAGTAGTAAAGCATATAGCCCTTGTCT 1572  
QY 576 ThrAsnSerLysValIleGlyGlyProGlyHisThrGlyValAsnLeuValTyrLeuGln 595  
DB 1573 TCAAGTGTCTCATTAATTAAGTTCACAGATTCACAGAGAAATTTACTATCTCTTAAA 1632  
QY 596 -----SerGlnGlyArgLeuGlnIleThrCysGlnThrProAsnSerThrGln 611  
DB 1633 GAATCTAGTAATTAATTCATTCCTAAATTAAAGTTACATTAATTAATTCAGACCTTGTACAA 1692  
QY 612 SerTyrPheIleArgLeuAspGlyTyrAlaThrAsnGlyAlaGlyAsnThrLeuProAsnIle 631  
DB 1693 CGATATCGGTAGAGATAGCGTATCTCTTACC-----ACTTACTTA 1734  
QY 632 SerLeuThrIleProGlyValIleGlyIleProGlyArgLeuAsnAsnThrPheSer 651  
DB 1735 CGACTTTTGTG-----CAAAATTCAAACATGATTTTCTT 1770  
QY 652 GlyThrAsnTyrAsn-----AsnLeuGlnTyrGlyAspPheGly 664  
DB 1771 GTCATCTACATTAATAAATACTATGATTAAGATGATGATTAATTAACATTAACAAATTTGAT 1830  
QY 665 TyrPheGlnPheProSerThrValThrLeuProLeuAsnArgAsnIlePhePheIlePhe 684  
DB 1831 CTCGCACTACTAATTAATTAATATGAGGCTTCTCGGTGATAGAGAACTTAATAATAGCA 1890  
QY 685 AsnAlaGalaAspValSerAsnSerIleLeuIleIleAspLysIleGlyPheIleProIle 704  
DB 1891 GCAGAAATCTTTCGTTCTTAATTAATAAATACTATATAGATAGATAGAAATTTATCCACGTA 1950  
RESULT 15  
US-07-973-320-3  
; Sequence 3, Application US/07973320  
; Patent No. 5286486  
; GENERAL INFORMATION:  
; APPLICANT: Payne, Jewel M.  
; TITLE OF INVENTION: No. 5286486el Bacillus thuringiensis Gene  
; TITLE OF INVENTION: Encoding a Coleopteran-Active Toxin  
; NUMBER OF SEQUENCES: 4  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: David R. Saliwanchik  
; STREET: 2421 N.W. 41st Street, Suite A-1  
; CITY: Gainesville  
; STATE: FL  
; COUNTRY: USA  
; ZIP: 32606  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/973,320  
; FILING DATE: 19921106  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/788,638  
; FILING DATE: 6-NOV-1991  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Saliwanchik, David R.

REGISTRATION NUMBER: 31,794  
REFERENCE/DOCKET NUMBER: MA68.C1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 904-375-8100  
TELEFAX: 904-372-5800  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 3414 base pairs  
TYPE: NUCLEIC ACID  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
ORIGINAL SOURCE:  
ORGANISM: Bacillus thuringiensis  
STRAIN: kumamotoensis  
INDIVIDUAL ISOLATE: HD867  
IMMEDIATE SOURCE:  
LIBRARY: Lamdegem (TW)-11 library of J.M. Fu  
CLONE: 867  
US-07-973-320-3  
Alignment Scores:  
Pred. No.: 1,18e-59 Length: 3414  
Score: 660.50 Matches: 223  
Percent Similarity: 43.1% Conservative: 112  
Best Local Similarity: 28.7% Mismatches: 286  
Query Match: 17.1% Indels: 157  
DB: 2 Gaps: 37  
US-10-783-417-2 (1-735) x US-07-973-320-3 (1-3414)  
QY 1 MetAsnGlnAsnAspAsnAsnGlnIleIleAspSerHisThrSerProTyr 20  
DB 1 ATGAATTTAAATTAATTAAGTGA--TATGAA-----GATAGT----- 36  
QY 21 PheProAsnArgAsnSerAsnAspSer--ArgTyrProTyrThrAsnAsnProAsnGln 39  
DB 37 -----AATAGAACATTAATTAATTTCTCAATTAATCTTACGCAAAAAGCATTAACACA 90  
QY 40 ProLeuGlnAsnThrAsnTyrIleGlyIleThrLeuAsnMetCysGlnGlyAsnThrGlnTyr 59  
DB 91 TCATTAAAGAAATATGAACCTACACAGCATTTTATCTATATCTATACGAG----- 135  
QY 60 GlyAspAsnPheGlnThrPheAlaSerAlaAspThrIleAlaValaSerAlaGlyThr 79  
DB 136 AGGGAACGCTGGAAGCACTGCTAGTGTATATACA-----GCTATTAAATCTGTAGTT 189  
QY 80 IleValSerGlyThrLeuLeuAlaGlyIleGlyGlyLeuThrSerIleSerGlyProIle 99  
DB 190 AGGTGTACAGGGGCTACACTAAGTCCGTAGGT-----GTCCAGGT----- 231  
QY 100 GlyIleIleGlyAlaIleIleIleSerPheGlyThrLeuIleThr--ValPheTyrPro 118  
DB 232 -----GCAAGTTTATCATTAACCTTAACCTGAATAATTAACGCGCTTTATAGGCA 282  
QY 119 AlaGlyGlnAspLysThrValTyrPheThrGlnPheIleLysMetGlyGlnIlePheVal 138  
DB 283 -----CACGATTAATAATTAATTTGGATGAATTTATGACAGAAAGTAAACACTTATT 333  
QY 139 AspThrProLeuThrGlnSerIleLysGlnLeuLysLeuGlnThrLeuGlnGlyPheArg 158  
DB 334 GAACAAAAAATAGAACATATATGACAGAAATAAGACATTCGCAATTTAGAGGATTAAGA 393  
QY 159 GlnIleLeuGlnSerTyrAsnThrAlaLeuAspAspTyrPargLysLeuLysArgLeuGln 178  
DB 394 AATAACTTAAGATATATCAACAGGCACTTGAAGATTGG-----CTGAAC 438  
QY 179 AlaProGlyLeuProProSerSerAlaLeuGlnGlnAlaIleLeuThrLeuLysIle--- 197  
DB 439 AATCTGTATGATCCA-----GCAACTATTAACAGAGTATGAT 477

QY 198 ArgPheGluAsnValHisAsnAspPheIleArgGluIleProGlyPheGluLeuGluThr 217  
 |||||  
 |||||  
 DB 478 CGTTTCGTATATTAGATCGTTTATTGAATCATATATGCGCATTTGCGTGGG 537  
 |||||  
 |||||  
 QY 218 TyrIleThrLeuLeuLeuProIleTyrAlaGlnAlaAlaAsnPheIleLeuLeuLeu 237  
 |||||  
 |||||  
 DB 538 TATGAATTCATCTTACTTAACAGTTTACGCAACAGCGCAACCTTCATCTAGCTTATTA 597  
 |||||  
 |||||  
 QY 238 GlnGlnGlyAlaGluLeuAlaAspGluTTPAsnAlaAspIleHisProSerGlnIleGlu 257  
 |||||  
 |||||  
 DB 598 AGGATTCCTACTCTTTATGAGATTAATG-----GAATTCCTCAGAACCAATTTGAG 651  
 |||||  
 |||||  
 QY 258 ProAsnAlaGlyThrSerAspAspTyrTyrIleLeuLeuLeuGluAsnIleProIleTyr 277  
 |||||  
 |||||  
 DB 652 GAAAT-----TATATCGTCAAAAGAAACATATTTCTGAATAT 690  
 |||||  
 |||||  
 QY 278 SerAspTyrCysAlaAsnThrTyrArgThrGlyLeuLeuAsnLeuArgAspGluProAsn 297  
 |||||  
 |||||  
 DB 691 TCTAACCATTCCTTATAGTGTATTAATAGTGTCTTACGAGATTGAAACGTTCCACTTAT 750  
 |||||  
 |||||  
 QY 298 MetIleTTPSerIlePheAsnAspTyrArgArgTyrMetThrIleThrValIleAspThr 317  
 |||||  
 |||||  
 DB 751 GAACAATGATTAATATATATATCGTTTCGTAGAAATGATTAATGATTAATAT 810  
 |||||  
 |||||  
 QY 318 IleSerGlnPheSerLeuTyrAspIleLeuArgTyrArgAspSerIleGlyGlyIleGlu 337  
 |||||  
 |||||  
 DB 811 GCGCGTATTCCTTATTTATGACCTCGAATGAT-----TCAATGGAA 855  
 |||||  
 |||||  
 QY 338 ValIleGlyIleIleAsnGluLeuThrArgGluIleTyrThrThrGluIleAsnPheAsp 357  
 |||||  
 |||||  
 DB 856 ACA-----AGTACGCGATTAAACGAGAGAGATGATTAACGATCAATTAAGCTTGCA 906  
 |||||  
 |||||  
 QY 358 ArgLeuProGlnLeuArgValGlnProAsnLeuAlaThrMetGluTyrAsnLeuThrArg 377  
 |||||  
 |||||  
 DB 907 ---ATTAGCAATCCAGGTATAGTCCAGTTTTCAGATGGAATAATCTCGGATTAAG 963  
 |||||  
 |||||  
 QY 378 AlaSerPheIleLeuPheSerPheLeuGluGlnPheIlePheTyrThrGluAsnThrAsn 397  
 |||||  
 |||||  
 DB 964 AACCA-----CACCTTGTTGATTTTATGATGAGCTTATATATATACA----- 1008  
 |||||  
 |||||  
 QY 398 PheGlyAsnArgLeuValGlyIleSerAsnArgAspAlaPro----- 411  
 |||||  
 |||||  
 DB 1009 -----TCAAAATATTAAGCATTTTCAACATGAGATTCAACACGATTTTATGAGAT 1062  
 |||||  
 |||||  
 QY 412 -----ThrTyrSerAsnThrIleThrGluThrLeuTyr 422  
 |||||  
 |||||  
 DB 1063 GCACATTAAGTTAGCTTTAAACAATCGAGCAATTCATTAATATATACAACAGCATATAT 1122  
 |||||  
 |||||  
 QY 423 GlyGluArgThrGly-----SerProThrThrIleArgProPheGluSerTyr 440  
 |||||  
 |||||  
 DB 1123 GGTAAACAAGTGGATATATTTATTCATCGAGGCGCATATTCATTTAGAGTAAATGATATCAT 1182  
 |||||  
 |||||  
 QY 441 LysVal-----SerIleValThrAspArgIleSerProValSerProIle 456  
 |||||  
 |||||  
 DB 1183 AGAACATTAGCAGCTCCATCAGTTGA-----GTTTATCCGTAT 1221  
 |||||  
 |||||  
 QY 457 GlnProHisPheIleIleAsnGlnIleGluLeuTyr-----Leu 469  
 |||||  
 |||||  
 DB 1222 ACTCAGATTAATGCTGCGAGAGTTGAGTTTACGGTGAATAAGGCAAGCATTAAT 1281  
 |||||  
 |||||  
 QY 470 AsnGlySerSerAsnAsnThrIleuLysTyrSerAlaGlyIleSerLeuSerAsnTyrGln 489  
 |||||  
 |||||  
 DB 1282 AGAGGATTAACAATAATGATCTGACGTATGAT----- 1314  
 |||||  
 |||||  
 QY 490 AsnThrThrPhePheGlnPheProArgLysAspCysAsnLeuValIleAspProGly 509  
 |||||  
 |||||  
 DB 1315 -----TCTATGATCAATTAAC-----CCAGAC 1338  
 |||||  
 |||||  
 QY 510 CysSerProAsnAsnAsnThrTyrSerHisIleLeuSerHis-----PheSerLeu 526  
 |||||  
 |||||  
 DB 1339 GGAGAACCAATACACAAAAATCACTCATGATTAATGATTAATGATTAATCAAGCTATATCTAAA 1398  
 |||||  
 |||||  
 QY 527 PheThrTyrSerTyrValIleGlyLeuGlnLeuGlnIleLeuAspThrGly----- 543

DB 1399 TCACTCCGCAATAT-----GATTAATCTACTATCCG 1431  
 |||||  
 |||||  
 QY 544 ValIleGlyTTPThrHisSerSerValAspArgTyrAsnAlaIleSerAspIleIle 563  
 |||||  
 |||||  
 DB 1432 ATCTTTCTTGACGATGAAGTCCGAGTATTAACAATTAAGATTAATCAACAAATC 1491  
 |||||  
 |||||  
 QY 564 ThrMetIleProAlaIleLeuGlyAsnAsnLeuAspThrAsnSerLysValIleGlyGly 583  
 |||||  
 |||||  
 DB 1492 ACMAAATTCAGCTCTATAAATGATTAACATGATGATGATCACTGATTCGAAAGG 1551  
 |||||  
 |||||  
 QY 584 ProGlyHisThrGlyIleAsnLeuValTyrLeuGlnSerGlnGlyArgLeuGluIleThr 603  
 |||||  
 |||||  
 DB 1552 CCGATTAACAGTGGAGATTTAGTTAAGAGAGGAGTAATGATTAATAGAGATATA 1611  
 |||||  
 |||||  
 QY 604 CysGluThrProAsnSer-----ThrGlnSerTyrPheIleArgLeuArgTyrAlaThr 621  
 |||||  
 |||||  
 DB 1612 AAGCTACCGTAACCTCACACTTCTCAAAATATCGTTAGAGCTTGATACGCCACT 1671  
 |||||  
 |||||  
 QY 622 AsnGlyAlaGly-----AsnThrLeuProAsnIleSerLeuThrIleProGlyValIle 639  
 |||||  
 |||||  
 DB 1672 AATGTTTTCGACATTCACCGTATATTAATGATTAATAATACGCT----- 1719  
 |||||  
 |||||  
 QY 640 GlyIleProProGlnArgLeuAsnAsnThrPheSerGlyThrAsnTyr-----AsnAsnLeu 658  
 |||||  
 |||||  
 DB 1720 -----CAAAAGAAAGTTTCAAAATACTGTAGAAACAATAGTGAAGAAAGATTA 1770  
 |||||  
 |||||  
 QY 659 GlnTyrGlyAspPheGlyTyrPheGlnPheProSerThrValThrLeuProLeuAsnArg 678  
 |||||  
 |||||  
 DB 1771 ACCATGTTTCATTTGGATATATAGAAATTTCTACACCATTCATTAATTCG--GATTAAG 1827  
 |||||  
 |||||  
 QY 679 AsnIleProPheIlePheAsnArgAlaAspValSer---AsnSerIleLeuIleIleAsp 697  
 |||||  
 |||||  
 DB 1828 CATCCAAAATACCTCTTATTAAGATTAATGAGTAAATTCATCAATTTATATATGAT 1887  
 |||||  
 |||||  
 QY 698 LysIleGluPheIleProIleThrSerSerMetHisGlnAsnArgGluLysGlnLysLeu 717  
 |||||  
 |||||  
 DB 1888 TCAATCGATTTATCCCTGATGATTAATATGAT-----GAAAAAGAAAAACATA 1938  
 |||||  
 |||||  
 QY 718 GluThrIleGlnThrLysIleAsnThrPhePheThrAsnHisThrLysThrLeu 735  
 |||||  
 |||||  
 DB 1939 GAAAAAGCACAGAAAGCCGTGAATACCTTGTTTACAGAGGAAAGAAATGCACATC 1992  
 |||||  
 |||||

Search completed: January 20, 2006, 22:34:14  
 Job time : 411 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: January 20, 2006, 16:46:39 ; Search time 7267 Seconds  
(without alignments)  
4732.146 Million cell updates/sec

Title: US-10-783-417-2

Perfect score: 3869  
Sequence: 1 MNQNNNDNNVEYIDSHTSPTV.....KLETIQKINTEFFTHNTKTL 735

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 41078325 seqs, 23393541228 residues

Total number of hits satisfying chosen parameters: 82156650

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODE=frame+ p2n.model -DEV=xlh  
-O=/cgn2.1/USPTO.spool/US10783417/runat.20012006.095010.23879/app.query.fasta\_1.903  
-DB=BSPT -QFMT=fastap -SUFFIX=p2n.rst -MINMATCH=0.1 -LOOPEXT=0 -LIST=45  
-UNITS=bits -START=1 -END=-1 -MATRIX=blomsum62 -TRANS=human40.cdi -LISF=45  
-DOCALIGN=200 -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=pco -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000  
-USER=US10783417@cgn 1 1 5315 @runat.20012006.095010.23879 -NCPU=6 -ICPU=3  
-NO MMAP -NEG SCORES=0 -WAIT -DSPBLOCK=100 -DEV TIMEOUT=120 -WARN TIMEOUT=30  
-THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7 -YGAPOP=10  
-YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

EST:\*  
1: gb\_esc1:\*  
2: gb\_esc2:\*  
3: gb\_esc3:\*  
4: gb\_hic:\*  
5: gb\_esc4:\*  
6: gb\_esc5:\*  
7: gb\_esc6:\*  
8: gb\_esc7:\*  
9: gb\_gsa1:\*  
10: gb\_gsa2:\*  
11: gb\_gsa3:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	147.5	3.8	6567	11	DQ031803 Homo sapi
2	124	3.2	13107	11	DQ035722 Homo sapi
3	121.5	3.1	3477	10	CL977084 OaIFCC029
4	121	3.1	1754	10	AG392933 Mus muscu
5	121	3.1	2032	4	CNS0A8VA BX819057 Arabidops
6	119.5	3.1	6777	10	CL967047 OaIFCC014
7	117.5	3.0	1986	10	AY416726 Mus muscu

8	117	3.0	2157	4	AK089255 Mus muscu
9	115	3.0	1178	1	AJ538005 AJ538005
10	115	3.0	2322	1	CL960568 OaIFCC004
11	115	3.0	2460	4	CNS09YP2 BX832541 Arabidops
12	114.5	3.0	820	7	CF951415 CF951415
13	112.5	2.9	1123	10	AG378595 Mus muscu
14	112.5	2.9	2841	10	CL974457 OaIFCC025
15	111	2.9	1143	4	CNS0A1JB BX829716 Arabidops
16	111	2.9	1516	10	AG396120 Mus muscu
17	110.5	2.9	1105	3	BM463012 BM463012
18	109.5	2.8	752	7	CO096813 CO096813
19	109.5	2.8	861	7	CO070923 CO070923
20	109	2.8	1878	2	BF528790 BF528790
21	108.5	2.8	949	7	CV674212 CV674212
22	108.5	2.8	956	4	AY122233 AY122233
23	108.5	2.8	2282	4	AK044734 AK044734
24	108.5	2.8	5001	4	CR860285 CR860285
25	107.5	2.8	2223	10	CL981417 CL981417
26	107.5	2.8	2439	10	CL982610 CL982610
27	107.5	2.8	2955	10	CL947058 CL947058
28	107	2.8	859	9	AZ548115 AZ548115
29	106.5	2.8	1541	4	BC015498 BC015498
30	106.5	2.8	1598	4	BC030976 BC030976
31	106.5	2.8	2355	10	AY413036 AY413036
32	106.5	2.8	2948	4	CNS0A4RI BX822972 Arabidops
33	106.5	2.8	4751	4	CR859511 CR859511
34	106	2.7	881	9	BH162188 BH162188
35	106	2.7	910	9	AZ546690 AZ546690
36	106	2.7	2520	10	AG280133 AG280133
37	106	2.7	3054	10	CL974410 CL974410
38	105.5	2.7	1092	10	CL065922 CL065922
39	105	2.7	622	7	CV176047 CV176047
40	105	2.7	697	7	CV533100 CV533100
41	105	2.7	747	3	BQ311235 BQ311235
42	105	2.7	16387	11	DQ038517 DQ038517
43	105	2.7	16425	11	DQ038516 DQ038516
44	104.5	2.7	458	1	AU006750 AU006750
45	104.5	2.7	872	6	CD792427 CD792427

#### ALIGNMENTS

RESULT 1  
LOCUS DQ031803 6567 bp DNA linear GSS 02-JUN-2005  
DEFINITION Homo sapiens HC2093 gene, VIRTUAL TRANSCRIPT, partial sequence,  
genomic survey sequence.  
ACCESSION DQ031803  
VERSION DQ031803.1 GI:66883012  
KEYWORDS GSS.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Homidae; Homo.  
REFERENCE 1 (bases 1 to 6567)  
Nielsen,R., Bustamante,C., Clark,A.G., Gianowski,S., Sackton,T.B.,  
Hubisz,M.J., Fledel-Alon,A., Tanenbaum,D.M., Civeello,D.,  
White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M.  
A Scan for Positively Selected Genes in the Genomes of Humans and  
Chimpanzees  
(er) PLOS Biol. 3 (6), E170 (2005)  
JOURNAL PUBLISHED 2 (bases 1 to 6567)  
15869325  
REFERENCE Nielsen,R., Bustamante,C., Clark,A.G., Gianowski,S., Sackton,T.B.,  
Hubisz,M.J., Fledel-Alon,A., Tanenbaum,D.M., Civeello,D.,  
White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M.  
Direct Submissio Submited (05-MAY-2005) Celera Genomics, 45 West Gude Drive,  
Rockville, MD 20850, USA  
COMMENT This sequence was made by sequencing genomic exons and ordering  
them based on alignment. Translation starts at the beginning of  
alignment.





Db 3460 AATTACATATATTTGGAGAGCTTCCACCATTTACATTTATATAGCTTTTTCCTCCGCA 3519  
 QY 534 GYLeuGlnLeuGlnIleLeuAspThrGlyValLeu---GlyThrPhHisSerVal 552  
 Db 3520 AGAAGTGAAGAAAGACTGGTCTTCAGTATCTTTCTTTTACAGATGAGTCAGTG 3579  
 QY 553 -----AepArgTyAsnAlaIleSerAspLysIle----- 562  
 Db 3580 CCGTTAGACACTCCACAAAATTGACTTATATCACTGATCTTACACTTTGATGCGTG 3639  
 QY 563 -----IleThrMetIleProAlaIleLysGlyAsn 572  
 Db 3640 AATAGAGCCCAAGTCTCTTCAGGTGGTATGTAAAGTATATAGTTTAAATTAATTCAT 3699  
 QY 573 AsnLeuAspThrAsnSer-----LysValIleGlnGlyProGlyHisThrGlyGly 589  
 Db 3700 GAACATGAACCTGACCTATATATATTAAGAATATATACAGA---TTTAAACTGAAGCC 3756  
 QY 590 AsnLeuValIleLeuGlnSerGlnIleArgLeuGlnIleThrCysGlnIleProAsnSer 609  
 Db 3757 AAACCTTTGACTGAGAA-----CCAGTCACG 3783  
 QY 610 ThrGlnSerTyPheIleArgLeuArgTyAlaThrAsn---GlyAlaGlyAsnThrLeu 628  
 Db 3784 ACC-----TACTCTATCCGTGATCTGCGTTCACCAAGTGGAAATGCAATCAATTT 3837  
 QY 629 ProAsnIle-----SerLeuThrIleProGlyValIle 639  
 Db 3838 AGTAAATGATGTAATAATTCAACCCCAAGATCAGTCCAGATGTCGTG 3885  
 RESULT 2  
 LOCUS DQ035722 13107 bp DNA linear GSS 02-JUN-2005  
 DEFINITION Homo sapiens DNAS gene, VIRUTAL TRANSCRIPT, partial sequence.  
 ACCESSION DQ035722  
 VERSION DQ035722.1 GI:66886931  
 KEYWORDS GSS.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
 Homiidae; Homo.  
 REFERENCE 1 (bases 1 to 13107)  
 AUTHORS Nielsen,R., Bustamante,C., Clark,A.G., Glanowski,S., Sackton,T.B.,  
 Hubisz,M.J., Fledel-Alon,A., Tanenbaum,D.M., Clevello,D.,  
 White,T.J., Snieky,J.J., Adams,M.D. and Cargill,M.  
 A Scan for Positively Selected Genes in the Genomes of Humans and  
 Chimpanzees  
 (er) PLOS Biol. 3 (6), E170 (2005)  
 JOURNAL PUBMED 15869325  
 REFERENCE 2 (bases 1 to 13107)  
 AUTHORS Nielsen,R., Bustamante,C., Clark,A.G., Glanowski,S., Sackton,T.B.,  
 Hubisz,M.J., Fledel-Alon,A., Tanenbaum,D.M., Clevello,D.,  
 White,T.J., Snieky,J.J., Adams,M.D. and Cargill,M.  
 Direct Submission  
 Submitted (05-MAY-2005) Celera Genomics, 45 West Gude Drive,  
 Rockville, MD 20850, USA  
 COMMENT This sequence was made by sequencing genomic exons and ordering  
 them based on alignment. Translation starts at the beginning of  
 alignment.  
 FEATURES  
 source  
 1..13107  
 /organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9606"  
 /chromosome="5"  
 <1..>13107  
 /gene="DNAS"  
 /locus\_tag="HC4910"  
 ORIGIN  
 Alignment Scores:

Pred. No.: 0.536 Length: 13107  
 Score: 124.00 Matches: 140  
 Percent Similarity: 35.2% Conservative: 86  
 Best Local Similarity: 21.8% Mismatches: 218  
 Query Match: 3.2% Indels: 198  
 Db: 11 Gaps: 35  
 US-10-783-417-2 (1-735) x DQ035722 (1-13107)  
 QY 97 GlyProIleGlyIleIleGlyAlaIleIleSerPheGlyThrLeuIleThrValPhe 116  
 Db 3676 GCACCTATTAGAGAAATCTTATGCCCCGCTTAAACATATAGACTTCTGATAGCA----- 3729  
 QY 117 TrpProIaGlyGlnAspLysThr-----ValIleThrGlnPheIle 131  
 Db 3730 -----AGGAAAGAGATAGCAAAAGTTGATACAGCTGCACTATGCTGGAGAAAGCTGCG 3783  
 QY 132 ---LysMetGlyGlnIlePheValAsp---ThrProLeuThrGlnSerIleLysGlnLeu 149  
 Db 3784 GCACGTCGCGGAGAGTCCAGAAATTAATAGTCTACCTGACGCCAGCTTCAAGAAAGAG 3843  
 QY 150 LysLeuGlnThrLeuGlnGlyPheArgGlnIleLeuGlnSerTyAsnThrAlaLeuAsp 169  
 Db 3844 CTTATATTAGTGTGTGAGAGTATCTCCCAAGATTGTCAACAGTTTATCTGACTATGAT 3903  
 QY 170 AspTrpArgLysLeuLysArgLeuGlnAlaProGlyLeuProSerSerAlaLeuGln 189  
 Db 3904 -----TTGAATGCTCAATGCTGACCGCTTAAAGCCAGAGAGCC---AGT 3948  
 QY 190 GlnAlaAlaLeuThrLeuLysIleArgPheGlnAsnValHisAsnAspPheIleArgGln 209  
 Db 3949 GACAGGCTTATCATGTTTTCAGATCAATTTGATTAATATCTATTCGGAATATACATCATAT 4008  
 QY 210 IleProGlyPheGlnLeu-----GluThrTyLeuThrLeuLeuProIle 225  
 Db 4009 ACTGAGAGAGAGAGAGCTTTTGGCCGTCAGCTACACAGTATCTTCAGCTTTGAAATTA 4068  
 QY 226 TyAlaGlnAlaAlaAsnPheHisLeuAsnLeuGlnGlnGlyAlaGlnLeuAlaAsp 245  
 Db 4069 AAGAGAGCA-----CTAAATCTTACACAAAATATATATCTG----- 4107  
 QY 246 GluTrpAsnAlaAspIleHisProSerGlnIleGluProAsnAlaGlyThrSerAspAsp 265  
 Db 4108 -----TACACAGCTTCATAGAA-----ACTGTAATATAGC 4137  
 QY 266 TyrTyLysLeuLeu-----LysGlnAsnIleProLysTySerAsnTyTyCys----- 281  
 Db 4138 TATTATGATATCTTTGTCAGAGGTGATTTGAAAATAATTAACATGAATCACTTAGAA 4197  
 QY 282 ---AlaSerThrTyArgThrGlyLeuLysAsnLeuArgAspGluProAsnMetLysTrp 300  
 Db 4198 TTCAGAAACGATGTCGAAAGCTTCCCGGCTTGAAGAGAC-----TGG 4242  
 QY 301 SerIlePheAsnAspTyArgArg-----308  
 Db 4243 CAGGCTTTTGGACCTGAAGAAATCATGATGATTTAGCGAGCTTCCCGCTGCGTG 4302  
 QY 309 ---TyrMetThr-----IleThrValLeuAsp 316  
 Db 4303 GAATCATGAGCAGATTAAGCCATGATGAGCGGCACTGGGAAAGATTAACACCTTCACC 4362  
 QY 317 ThrIleSer-----GlnPheSerLeuTyAspIleLys----- 327  
 Db 4363 GGGCAACAGTGTGATGTGGGAAATGAAGAACTTTAAGTTAAGAAATTCATGAGAGCACT 4422  
 QY 328 -----ArgTyArgAspSerIleGlyGlyIleGlyValIleGlyLys----- 342  
 Db 4423 CTTTGAATATTAAGAGAAATTAAGAGCACTGTGATAGTGTGCGGTGAAGAGAGAGAC 4482  
 QY 343 -----AsnGln----- 344  
 Db 4483 ATTGAGCAAAAGCTGAAGCAAGTGAATTAATGATGGAACAATTAATCACTTCGCGC 4542

```

Oy 345 -----LeuThrArgGluIleTyrThrThrGluIleAsnPro 356
Db 4543 AGCTTTAAACCCGTGAGAGCTCCTCTTGAAGAGAGACAGTACCTCGGAATACCTGCC 4602
Oy 357 AspaArgLeuProGlnLeuArgValGlnProAsnLeuAlaThiMetGlnTyrAsnLeuThr 376
Db 4603 AACATGAGAGACAGCTGATGTTGCTGGAGATCCCTACGACAGCAACAGGACAAATG--- 4659
Oy 377 ArgAlaSerPheLeuPheSerPheLeuGlnGlnPheLeuPheTyrThrGluAsn--- 395
Db 4660 -----CCATTCAA-----GCCAGATCAAAAATGGTGTCAGTACCTTCCAACTCA 4707
Oy 396 ThrAsnPheGlyAsnArgLeuValGlyIleSerAsn-----ArgAspAlaPro 411
Db 4708 ACAGACATCATCAGAGCTGATGACGGTGCAAAAACCTGTGATTATTAAGAACCTGTC 4767
Oy 412 ThrTyrSerAsnThrIleThrGlnThrLeuTyrGlyGlnArgThrGlySerProThrThr 431
Db 4768 TTTGTGGAGAGAGACATTCGCCAAGCAGCTGCCCAAGAA-----GCC 4809
Oy 432 LysThrIleArgProPheGluSerTyrLysValSerIleValThrAspArgGlnSerPro 451
Db 4810 AAGCGCTTTCTTAACATAGATAAATCTTGGGTGAAGATATCATCTCGGACATGAATG 4869
Oy 452 Pro-----ValSerProIleGlnProHisPhe 460
Db 4870 CCCAGTATGATCCAGCTGCTGTGTGGAGATGAGACCTCGGGCAGCTGTACCAACAC--- 4926
Oy 461 IleIleAsnGlnIleGlnLeuTyrLeuAsnGlySerSerAsnAsnThrLeuLysTyrSer 480
Db 4927 TTGCTGACCAAGTTGGAATA----- 4947
Oy 481 AlaGlyLysSerLeuSerAsnTyrGlnAsnThrThrPhePheGlnPheProArgLysLys 500
Db 4948 TGCCGAAATATCCCTTACTGCGTACTTGAGAAAAAACACATGCTTCTCTCGG----- 5001
Oy 501 AspCysAsnLeuValIleAspProGly-----CysSerProAsn 513
Db 5002 ---TTTTCTTCTGCTCAGATCTGCGCTTCTAGAGATTCTGGGACAGCGTCGAGACTCC 5058
Oy 514 PheAsnAsnTyrSerHisIleLeuSerHisPheSerLeuPheThrTyrSerTyrValIle 533
Db 5059 CACACTATACAGGCCCATTTGCTGAATGTGTGAC----- 5094
Oy 534 GlyLeuGlnLeuGlnIleLeuAspThrGlyValLeuGlyTyrThrHisSerSerValAsp 553
Db 5095 -----AACATTAAATCTGTCAAG 5112
Oy 554 ArgTyrAsnAlaIleSerAspLysIleIleThrMetIleProAlaIleLysGlyAsnAsn 573
Db 5113 TTCCACGAAAGATCATGATCGAAATTCGTCA---ATTCTCTCAAGAGGGTGAGACG 5169
Oy 574 LeuAspThrAsnSerLysValIleGlnGlyProGlyHisThrGlyGlyAsnLeu---Val 592
Db 5170 ATTGAA-----TTGATTAACCTGTCACTGACAGAGGCGCATGTGAAGTT 5214
Oy 593 TyrLeuGlnSerGln---GlyArgLeuGlnIle-----ThrCysGlnThrProAsn 609
Db 5215 TGCGTTAATCTCTTTTGAAGAAATCTCAGTCTCATTCGATCTTGTAATGCGCAGGCA 5274
Oy 609 eTThrGlnSerTyrPheIleArgLeuArgTyrAlaThrAsnGlyAlaGlyAsnThrLeuP 629
Db 5275 GCCGCAAA-----TATTCAGAAACAGGTTTCCAACCTAAGAAATTTCTTCATCTTC 5328
Oy 629 ro 629
Db 5329 CC 5330

```

```

RESULT 3
LOCUS CL977084 3477 bp DNA linear GSS 21-SEP-2004
DEFINITION OsLFC029775 Oryza sativa Expressed Sequence Tag (EST)
cDNA clone, partial genomic sequence.

```

```

ACCESSION CL977084
VERSION CL977084.1 GI:52408678
KEYWORDS GSS.
SOURCE Oryza sativa (indica cultivar-group)
ORGANISM Oryza sativa (indica cultivar-group)
Bukeriyota, Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzaceae; Oryza.
1 (bases 1 to 3477)
Me, L., Wang, J., Chen, C., Liu, X., Su, N., Li, L., Wang, X., Cao, M.,
Jiao, Y., Sun, N., Zhang, X., Bao, J., Sun, D., Zhao, H., Yuan, L.,
Wong, G.K.S., Deng, X.W. and Wang, J.
An analysis of transcriptional regulation of the rice genome and
its comparison to Arabidopsis
Unpublished (2004)
JOURNAL
COMMENT
Contact: Chen Chen
Department of Bioinformatic
Beijing Institute of Genomics
Chinese Academy of Sciences, Beijing 101300, China
Tel: 86-10-80481559
Fax: 86-10-80488676
Email: chenchen@genomics.org.cn
Rice genomic sequence.
Classes: exon-trapped.
Location/Qualifiers
source 1..3477
/organism="Oryza sativa (indica cultivar-group)"
/mol_type="genomic DNA"
/db_xref="taxon:39946"
/clone_lib="Oryza sativa Expressed Library"
/note="Oryza sativa exon trapped genomic sequences "

```

```

ORIGIN
Alignment Scores:
Pred. No.: 0.122 Length: 3477
Score: 121.50 Matches: 144
Percent Similarity: 31.84 Conserved: 79
Best Local Similarity: 20.54 Mismatches: 243
Query Match: 3.14 Indels: 235
DB: 10 Gaps: 31
US-10-783-417-2 (1-735) x CL977084 (1-3477)
Oy 132 LysMetGlyIleIlePheValAsp---ThrProLeuThrGlnSerIleLys----- 147
Db 1105 AAAATGAACGAACCTTACTACAGCAAAACCAATTAACGGTTCATCCCAAGAAATA 1164
Oy 148 -----GlnLeuLysLeuGlnThrLeuGlnGlyPheArg 158
Db 1165 GGAATGCTAGCCCAACCTGCAAGTTTGCAGCTAAGCAACACACCTTGAGCGG----- 1218
Oy 159 GlnIleLeuGlnSerTyrAsnThrAlaLeuAspAspTPrArgLysLeuLysArgLeuGln 178
Db 1219 -----GAGATACCCCAACGCTCTAGCAAAATCTGACCAACCTGACTTGA 1266
Oy 179 AlaProGlyLeuProProSerSerAlaLeuGlnGlnAlaAlaLeuThrLeu---LysIle 197
Db 1267 CATATGATATGAACTGTCGGGGCTATATCCCAAAACCTGACAGCTACCAAGAG 1326
Oy 198 ArgPheGluAsnVal---HisAsnAspPheIleArgGlnIleProGlyPheGlnLeuGln 216
Db 1327 CAATTACTTAAGCTTACGAAAAACAACCTTACGCGAAATCCCGGTATTCAT 1386
Oy 217 ThrTyrLysThrLeuLeuLeuProLeyTyrAlaGlnAlaAlaAsnPheHisLeuAsnLeu 236
Db 1387 CTAAACAAAGTGAAAAAATTTACTATACCAAAACAA----- 1425
Oy 237 LeuGlnGlnGlyAlaGlnLeuAlaAspGluTPrAsnAlaAspIleHisProSerGlnIle 256
Db 1425 ----- 1425
Oy 257 GluProAsnAlaGlyThrSerAspAspTyrTyrLysLeuLeuLysGlnAsnIleProLys 276
Db 1425 ----- 1425

```

Db 1426 -----GTCAAGGGTTCCATCCCAAA 1446  
 QY 277 TyrSerAsnTyrCysAlaAsnThrTyrArgThrGlyLeuIleuAsnLeuArgAspGluPro 296  
 Db 1447 GAAATGGAATGTTCCCACTGACGCTTAAAGTCTAGGCAACAATACCTTGAACGA 1506  
 QY 297 AsnMetLeuTrpSerLeuPheAsnAspTyrArgArgTyrMetThrIleThrValLeuAsp 316  
 Db 1507 GAGATCCCAACAACCTATCAAAAT-----CTAACCAACCTAAGAT 1545  
 QY 317 ThrIleSerGlnPheSerLeuTyrAspIleuAspTyrArgAspSerIleGlyIle 336  
 Db 1546 ACTTTG-----TCCCTAATGGGAT----- 1563  
 QY 337 GluValIleGlyIleLeuAsnGluLeuThrArgGluIleTyrThrThrGluIleAsnPhe 356  
 Db 1564 -----AATGAACCTTGGGGCATATACCCCAAAACTGTCACGCTC 1605  
 QY 357 AspArgLeuProGlnLeuArgValGlnProAsnLeuAlaThrMetGluTyr----- 373  
 Db 1606 ACCAAGATGCAATACCTTAGCTTACAGACAAACAATTAAGCGAAATCCCTGGGTGT 1665  
 QY 374 -----AsnLeuThrArgAlaSerPheLeuPheSerPheLeuGlu----- 387  
 Db 1666 TTATCAATCTAACAATAATGAA-----AACTTACTATACCAAAACCAAGTCACGGGT 1722  
 QY 388 -----GlnPheIlePheTyrThrGluAsn 395  
 Db 1723 TCCATCCCAAGAAATAGGAATGCTTCCCAACCTGCAAGTTTGTGACCTAACCAACAC 1782  
 QY 396 ThrAsnPheGlyAsnArgLeuValGlyIleSerAsnArgAspAlaProThrTyrSerAsn 415  
 Db 1783 ACCTTAGCGGGAGATCCACAGCTCATCAAAAT-----CTAACCAAC 1827  
 QY 416 ThrIleThrGluThrLeuTyrGlyIleArgThrGlySerProThrThrIleArg 435  
 Db 1828 CTAGCATATTTGCTCATGGGTAAATGAACCTTGGGGCTATACCCCA----- 1878  
 QY 436 ProPheGluSerTyrIleValSerIleValThrAspArgIleSerProProValSerPro 455  
 Db 1879 -----AACTCTGCATGCTCACCAAGATCCAAATATCTTGAATCTTACAGAT 1923  
 QY 456 IleGlnProHisPheIleIleAsnGln-----IleGluLeuTyrLeuAsnGlySer 472  
 Db 1924 TCTCTTCTCGGAATTCGAAATCTCAGCGGCATACGACACTTGGCTGTGAATATAC 1983  
 QY 473 Ser-----AsnAsnThrLeuTyrTyrSerAlaGlyIleSerLeuSerAsnTyrGln 489  
 Db 1984 TCATTTTCAGACATTTGCTCGCAATATGATGATGGGTGGCAGACTCAAAACTTTCAGT 2043  
 QY 490 -----AsnThrThrPhePheGlnPheProArgLys---LysAspCys----- 502  
 Db 2044 ATTGTGGGAATGCTTCGATGAGCCCATTCACAGAAAGCTTGAAGACATGTACAGCTTG 2103  
 QY 503 -----AsnLeuValIleAspProGlyCysSerProAsnPheAsnAsn 516  
 Db 2104 GTTAAACCTTCCGTGACAAATTAACCTGCTTACGGAGACATATCTGAACATTTTGAAGTG 2163  
 QY 517 TyrSerHisIleLeuSer-----HisPheSerLeuPheThrTyr 529  
 Db 2164 TATCCGCACTCAAAATCGGTGACGCTTGTCAATACATAGTTCTTGGGCGAGATCCACCA 2223  
 QY 530 SerTyrValIleGlyLeuGlnLeuGlnIleLeuAsp-----ThrGly 543  
 Db 2224 AATGGGTGCTGATGCTCCAGCTTGAAGAAATGATTTTCACAAAACATGATCACAGGT 2283  
 QY 544 ValLeuGlyTyrPheHisSerValAspArgTyrAsnAlaIleSerAspIleIle 563  
 Db 2284 TTACTAGACTTATCAT-----AATATATCATGTGTGG----- 2319  
 QY 564 ThrMetIleProAlaIleuGlyAsn-----AsnLeuAspThr 576  
 Db 2320 -----ATTCCAGCAGATTTGGAATTCGAAAGCTATATAGATTAATTTATCATCTC 2373

QY 577 AsnSerLeuValIleGluGlyProGlyHisThrGlyIleAsnLeuValTyrLeuGlnSer 596  
 Db 2374 AAT-----CAGCTTCAGATATTTTACCGACGACAGCTAGGAAGTTAAGCAT 2421  
 QY 597 GlnGlyArgLeuGluIleThrCysGluThrProAsnSerThrGlnSerTyrPheIleArg 616  
 Db 2422 TTAGGGTACCTTGAATGTATCAGAAAC----- 2448  
 QY 617 LeuArgTyrAlaThrAsnGlyAlaGlyAsnThrLeuProAsnIleSerLeuThrIlePro 636  
 Db 2449 -----AATTGAGTGACCAATACCC 2469  
 QY 637 GlyValIleGly-----IleProProGln 644  
 Db 2470 GATGAACCTAGGGAGCATGCTACGATTCAGAGAAAGTCTTCATCATCGTAGCGCCG 2529  
 QY 645 ArgLeuAsnAsnThrPheSerGlyThrAsnTyrAsnAsnLeuGlnTyrGlyAspPheGly 664  
 Db 2530 ATGTGAGCTTAAACAGTTCTTGATGTGCTTACCAACACTTGGA----- 2574  
 QY 665 TyrPheGlnPheProSerThrValThrLeuProLeuAsnArgAsnIlePro---PheIle 683  
 Db 2575 ---GGACTCTCCAGCAGACATCTTCAA---AATGATCAATTAAGCTGTATT 2628  
 QY 684 PheAsnArgAlaAspValSerAsnSer----- 692  
 Db 2629 CACATTAAGATTATAGCATCTGTGCACCATTAAGCCAAAGCTAAGTTTCTACTGCCA 2688  
 QY 693 -----IleLeuIleIleAspIleGluPheIleProIleThrSerSerMet 708  
 Db 2689 ATTTGTTCTGTGTGTGATGTGATCTTGGGTACAAATATGTATGATCAAAAGCTTGT 2748  
 QY 709 HisGlnAsnArgGluIleGlnIleuGlnIleGlnThrIleAsnThrPhePhe 728  
 Db 2749 CATTAACAGAGAAACCAACAGCAAGCTGATGCTATTAAGCTTGCAAGAACATGTT 2808  
 QY 729 Thr 729  
 Db 2809 TCT 2811  
 RESULT 4  
 AG392993/c 1754 bp DNA linear GSS 21-DEC-2004  
 LOCUS Mus musculus molossinus DNA, clone:MSMg01-211C07.T0, genomic survey  
 DEFINITION  
 ACCESSION AG392993 1754 bp DNA linear GSS 21-DEC-2004  
 VERSION AG392993  
 KEYWORDS  
 SOURCE Mus musculus molossinus (Japanese wild mouse)  
 ORGANISM Mus musculus molossinus  
 Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;  
 Sciurognathi; Muridae; Murinae; Mus.  
 REFERENCE  
 AUTHORS 1 Abe, K., Noguchi, H., Tagawa, K., Yuzurika, M., Toyoda, A., Kojima, T.,  
 Ezawa, K., Saitou, N., Hattori, M., Sakaki, Y., Moriwaki, K. and  
 Shirotani, T.  
 TITLE Contribution of Asian mouse subspecies Mus musculus molossinus to  
 genomic constitution of strain C57BL/6J, as defined by BAC-end  
 sequence-SNP analysis  
 JOURNAL Genome Res. 14 (12), 2439-2447 (2004)  
 PUBMED 15574823  
 REFERENCE 2 (bases 1 to 1754)  
 AUTHORS Hattori, M., Toyoda, A., Noguchi, H., Kojima, T. and Sakaki, Y.  
 DIRECT SUBMISSION  
 TITLE Submitted (17-NOV-2003) Masahira Hattori, The Institute of Physical  
 and Chemical Research (RIKEN), Genomic Sciences Center (GSC),  
 1-7-22 Suenho-chou, Tsukuba, Ibaraki, Japan, 305-8565, Japan  
 (E-mail:hattori@gsc.riken.jp, URL: http://hgp.gsc.riken.go.jp/  
 Tel:81-45-503-9111, Fax:81-45-503-9170)  
 COMMENT Clones are derived from the mouse BAC library MSMg01. For BAC  
 library availability, please contact Kuniya Abe (abe@rtc.riken.jp).



FEATURES  
source  
Location/Qualifiers  
1. 2032  
/organism="Arabidopsis thaliana"  
/mol\_type="mRNA"  
/db\_xref="taxon:3702"  
/clone="GSLTFB40ZB03"  
/tissue\_type="flowers and buds"  
/ecotype="Col-0"  
/plasmid="pCMVSPORT\_6"  
1. 2032  
/gene="At2g35230"  
/gene="At2g35230"

ORIGIN

Alignment Scores:  
Pred. No.: 0.0582 length: 2032  
Score: 121.00 Matches: 95  
Percent Similarity: 35.2% Conservative: 72  
Best Local Similarity: 20.0% Mismatches: 168  
Query Match: 3.1% Indels: 139  
Gaps: 23

US-10-783-417-2 (1-735) x CNS0A8VH (1-2032)

```

Qy 306 TyrAgaGgTymeThrIleThrValLeuAspThrIleSerGlnPheSerLeuTyraP 325
    |||
Db 480 TACTATAGGATATTGATTTTGTGAGATGATAGG----- 515
Qy 336 IleLysArgTyrArgAspSerIleGly-----GlyIleGluValLysGlyIle 341
    |||
Db 516 ---CCTAGACAAATATGATCATTTGGGTGTGATAGGATGGAAAAATATCAGA----- 566
Qy 342 LysAenGluLeuThrArgGluIleTyThrThrGluIleAsnPheAspArgLeuProGln 361
    |||
Db 567 AAGAGCTCTTCATCATCAATGCACTTTCGTCGACAGTACTAGCAACGGTCCCTCCACAG 626
Qy 362 LeuArGValGlnProAsnLeuAlaThrMetGluTyAsnLeuThrArgAlaSerPheLys 381
    |||
Db 627 CTTCAACACAGCCTCAGGTT-----TATACATTAAGCAAGAACACATTACAA 674
Qy 382 LeuPheSerPheLeuGlnPheIlePheTyThrGluAsnThrAsnPheGlyAsnArg 401
    |||
Db 675 ---AGTATGTTTACAGCAG----- 689
Qy 402 LeuValGlyIleSerAsnArgAspAla-----ProThrTyrSerAsnThrIleThr 418
    |||
Db 690 CTAAAGGGGTTCTCCACAGCTGAAGCCTTCTCGGCTCTCCAGAACACTCGCTA--- 746
Qy 419 GluThrLeuTyGlyGluArgThrGlySerProThrThrIleArgProPheGlu 438
    |||
Db 747 -----AGGCTTCAGATACACGGTTGCAGCGGATCAGACCA----- 782
Qy 439 SerTyLysValSerIleValThr-----AspArgGlnSerProProValSerProIle 456
    |||
Db 783 -----TCCCCCTTAAGCATTAACACCGGCGCGGTTCTCTCCCTTCAG 830
Qy 457 GlnProHisPheIleIleAsnGlnIle-----GluLeuTyLysLeuAsnGly 471
    |||
Db 831 GCTCCGACAGCTCATCTCAGTTGGCTAGACAGCCTCCACACACGACCTTCCAC 890
Qy 472 SerSerAsnAsnThrLeuTySerIleGlyIleSerLeuSerAsnTyThrAsnThr 491
    |||
Db 891 AAACACACACACCAATGATGGGTACAGAGGACCAAGTTTGGTGAATACAG----- 944
Qy 492 ThrPheGlnPheProArgLysLysAspCys-----AsnLeuValIle 506
    |||
Db 945 -----CTGAGTTTCCCGTCCAGATATAGGTTATCTTCAAGTTAACTTGGAGATT 998
Qy 507 AspProGlyCysSerProAsnPheAsnAsnTySerHisIleLeuSerHisPheSerLeu 526
    |||
Db 999 CAG-----GACCAATGCTAACCAATGACGACGAGTCAATGACGAC----- 1040
Qy 527 PheThrTyrSerTyValIleGlyLeuGlnLeuGlnIleLeuAspThrGlyValLeuGly 546
    |||

```

```

Db 1041 -----GGCCTTATATCCAGAGTCATGACAGCGCATATG--- 1076
Qy 547 TrpThrHisSerSerValAspArgTyRsnAlaIleSerAspLysIleIleThrMetIle 566
Db 1076 ----- 1076
Qy 567 ProAlaIleLysGlyAsnAsn-----LeuAspThrAsnSerLysValIleGlu 582
    |||
Db 1077 -----TGCACAGGTATATGACAGACGACCATATATGCCAGGTATATGACAGCGGCATAT 1130
Qy 583 GlyProGlyHisThrGlyLysLeuValTyLysLeuGlnSerGlnGlyArgLeuGluIle 602
    |||
Db 1131 ATCCAGAGTCATGACAGCGGTATATATGACAGTCATCTCAGTCACTCAACT 1190
Qy 603 ThrCysGluThrProAsnSerThrGlnSerTyRpnIleArgLeuAspArgTyRatThrAsn 622
    |||
Db 1191 CAACCTCA-----CCTCAACCTCAGACATATGATGCTGACCGGACCTGCAATGAC 1247
Qy 623 GlyAlaGlyAsnThrLeuProAsnIleSerLeuThrIleProGlyValIle----- 639
    |||
Db 1248 ATGACAGGCGCCTTCACACTTAACCATATCTACACACCGCGGTATGTTCTTACGCCA 1307
Qy 640 -----GlyIleProProGlnArgLeuAsnAsnThrPheSerGlyThr----- 653
    |||
Db 1308 GTGCCTCATATCTTCTCTCTCGTTCAATGCTCCGTACCTGTGACGCTACTCAG 1367
Qy 654 -----AsnTyRsnAsnLeuGlnTyRglAspPhe 663
    |||
Db 1368 CCTCTCCCATGTTGACATGATATGATGATGATTTCTTCTTCATATTAACGCTTT 1427
Qy 664 GlyTyRpnGlnPheProSerThrValThrLeu-----ProLeuAsnArgAsnIleProPhe 682
    |||
Db 1428 GACACACGTGACATCACTATCCATCCATTTCTTCAACATCTCTTACCGGTTAACCGGAAT 1487
Qy 683 IlePheAsn-ArgAlaAspValSerAsn-----SerIleLeuIleI 696
    |||
Db 1488 ATGTTCTCTCCAGATACCGCTTATCCCATGCTGACACAGAGTTCAAGTATCTTC----- 1542
Qy 696 eAspLysIleGluPheIleProIleThrSerSerMetHisGlnAsnArgGluGlyIle 716
    |||
Db 1543 -----AACCACTTACCCCAACCTTCTATTTCAAAATGCTCAACAGAGAGT 1592
Qy 716 eLeuGluThrIleGlnThrLysIleAsnThrPhePheThr 729
    |||
Db 1593 CTTGACCCCGGTGACAGTCCAAATCAAGGTCTCTCTCAGC 1632

```

RESULT 6  
CL967047  
LOCUS  
DEFINITION  
Oryza sativa (indica cultivar-group) genomic, genomic survey sequence.  
ACCESSION  
CL967047  
VERSION  
CL967047.1 GI:52388742  
KEYWORDS  
SOURCE  
ORGANISM  
Oryza sativa (indica cultivar-group)  
Oryza sativa (indica cultivar-group)  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
Ehrhartoideae; Oryzoideae; Oryza.  
REFERENCE  
1 (bases 1 to 6777)  
Ma, L., Wang, J., Chen, C., Liu, X., Su, N., Li, L., Wang, X., Cao, M.,  
Jiao, Y., Sun, N., Zhang, X., Bao, J., Sun, D., Zhao, H., Yuan, L.,  
Wong, G.K.S., Deng, X.W. and Wang, J.  
An analysis of transcriptional regulation of the rice genome and  
its comparison to Arabidopsis  
Unpublished (2004)  
JOURNAL  
COMMENT  
Contact: Chen Chen  
Department of Bioinformatic  
Beijing Institute of Genomics  
Chinese Academy of Sciences, Beijing 101300, China  
Tel: 86-10-80481559  
Fax: 86-10-80488676  
Email: chenchen@genomics.org.cn

Rice genomic sequence.  
 Class: exon-trapped.  
 Location/Qualifiers  
 source  
 1..6777  
 /organism="Oryza sativa (indica cultivar-group)"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:39946"  
 /clone\_lib="Oryza sativa Expressed Library"  
 /note="Oryza sativa exon trapped genomic sequences"

## ALIGNMENT SCORES:

Pred. No.:	0.615	Length:	6777
Score:	119.50	Matches:	145
Percent Similarity:	34.1%	Conservative:	85
Best Local Similarity:	21.5%	Mismatches:	209
Query Match:	3.1%	Indels:	238
DB:	10	Gaps:	36

US-10-783-417-2 (1-735) x CL967047 (1-6777)

```

QY 151 LeuGlnThrLeuGluGlyPheArgGlnIleLeuGlnSerTyrrAsnThrAlaLeuAspAsp 170
   ::::: ||||| ||||| ::::: ||||| |||||
DB 427 ATATCATGCTTACTAGATTTCAGGCAAGCATCATGATCCACGGAGCGCTTAACAAGT 486
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 171 Trp-----ArgLysLeu 174
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 487 TGGACACGCGCGCTTCCCATTCGACGTGAAGGCGTCAGTTGTAGCCGAGGACGCTC 546
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 175 LysArgLeuGlnAla-----ProGlyLeuProProSerSerAlaLeuGln 189
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 547 GCGCAGTCAAGCGCGCTCAACCTTACACGAAAGATTATCGGGTTCCATTTCTCCTCT 606
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 190 GlnAlaAlaLeuThrLeuLysIleArgPheGlnAsnValHisAsnAspPheIleArgGlu 209
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 607 GTTGAAGACCTGACATCTCTTATACATGACCTGTCCATATATACCTGTCGGGAG 666
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 210 IleProGlyPheGlnLeuGlnThrTyrrLysThrLeuLeuProIleTyrrAlaGlnAla 229
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 667 ATGCTCT----- 672
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 220 AlaAsnPheHisLeuAsnLeuGlnGlnGlyAlaGlnLeuAlaAspGluThrAsnAla 249
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 673 -----CATCTTAACAATCTGCAAGAGATGACGCTCTTAATCTGAGCTACAATCTC 723
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 250 -----AspIleHisPro----- 253
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 724 TTAGATGAATTAATTCACAGACGCTCACCAATGTTCCAACTACAGCAATTAATCTTA 783
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 254 -----SerGlnIleGluProAsnAlaGly-----ThrSerAspAsp 265
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 784 AGATATTAATTAACCTTCGAAAGTCGTATTCCTCCGAAATAGTCTCCTTAACCAACTGTT 843
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 266 TyrTyrrLysLeuLeuLysGlnIle-----ProLysTyrrSerAsnTyrr 280
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 844 TACATGTCATGTCACATTAATATCTCTCTGCAATTAATCCCAACCTCGGTAACTG 903
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 281 CysAlaAsnThrTyrrArgThrGlyLeuLysAsn-----LeuArgAspGlu 295
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 904 ACCTGCGTAAGTAATCTCGCTCTTACACAAATCATCTTAAGAGAAAGATCCCAATGA 963
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 296 -----ProAsnMetLysTrpSerIlePheAsnAspTyrrArgTrpMetThr 311
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 964 ATTGGGAAACTTCCCAATATCAGCATCGTAGTCTTAGAGAT-----AACATGTTGCA 1017
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 312 IleThrValLeuAspThrIleSerGlnPheSerLeuTyrrAspIleLysArgTyrrArgAsp 331
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1018 GGTAGTATACCGAAACCTTCT-----AATTCCTTAACGGGG-----AATTCCT 1047
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 332 SerIleGlyGlyIleGluValLysGlyIleLysAsnGluLeuThrArgGluIleTyrrThr 351
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1048 TCTATTACAGGAATAGAAATTACGTGCA-----AATCTTCTTAACGGGG-----ACT 1092
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

```

QY 352 ThrGluIleAsnPhe---AspArgLeuProGlnLeuAlaGlnProAsnLeuAlaThr 370
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1093 CTACCTATAAATCATGGGTATATCTCCCTAACTCAAA-----TCTCTCTTC 1140
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 371 MetGluTyrrAsnLeuThrArg-----AlaSerPheLysLeuPheSerPheLeu 386
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1141 TTGGGCCCAACATCTTGAGGGGTCAATCCAGGCTCTTATAGGCAAGCTTCAACGCTTA 1200
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 387 GlnGlnPheIlePheTyrrThrGluAsn-----ThrAsnGlyLysn 400
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1201 AAACAAGTAGTTCACAAACCAACAATTTCATGCGCAGATTCCTCACTACTTGGAA--- 1257
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 401 ArgLeuValGlyIleSerAsnArgAspAlaProThrTyrrSerAsnThrIleThrGluThr 420
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1258 AAGCTTTTGGACCTCCAGACTAATCTTCAGCA-----AACATGCTTAAGCA 1308
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 421 LeuTyrrGlyLysArgThrGlySerProThrThrLysThrIleArgProPheGluSerTyrr 440
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1309 AAAGCACTGAAGGCTGGGGA-----TTTCTGATGCACTGAGAAACTACAGTGTCTTA 1362
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 441 LysValSerIleValThrAspArgGlnSerProProValSerProIleGlnProHisPhe 460
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1363 CAATCTTGAATTAATCAGAAATCACTACAGGGAATTAATCA----- 1407
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 461 IleIleAsnGlnIleGluLeuTyrrLeuAsnGlySerSerAsnAsnThrLeuLysTyrr 479
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1408 -----AATTCATTC-----GGTACCTATCGAACACTCTTGGTATCTC 1446
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 480 ---SerAlaGlyLysLeuSerSerAsnTyrrGlnAsnThrThrPheGlnPheProArg 498
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1447 ATGTTGGGTGAATAATGAATCTATAGC----- 1473
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 499 LysLysAspCysAsnLeuValIleAspProGlyCysSerProAsnPheAsnAsnTyrrSer 518
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1474 -----ATAGTCCCTCATGATCAATAGGGAACCTTCTGGTATGTT 1512
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 519 HisIleLeuSerHisPheSerLeuPheThrTyrrSerTyrrValIleGlyLeuGln----- 536
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1513 CAACTGACATTAGATGCAACAGCTTAACAGGTTGATGACAGTCCATGATTTGAATCTT 1572
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 537 -----LeuGlnIleLeuAsp-----ThrGlyValLeuGlyLysProThrHisSerSe 551
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1573 AAATATCTTGAATTAATTAATCTTGGAAAGAACAGATTCATCGGCGCTTAATTCATAT-TC 1631
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 551 rValAspArgTyrrAsnAlaIleSerAspLysIleIleThrMetIleProAlaIleLysGlu 571
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1632 TATTGGCAGACTTACACGGCTGACAGAG----- 1659
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 571 YAsnAsnLeuAspThrAsnSerLysValIleGluGlyProGlyHisThrGlyLysnLe 591
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1659 ----- 1659
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 591 uValTyrrLeuGlnSer-----GlnGlyArgLeuGlnIleThrCysGluThrProAs 608
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1660 -CTTATCTGGAAGAACATGCAATTGGAAGAACCAATCCAAATTTCT----- 1704
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 608 nSerThrGlnSerTyrrPheIleArgLeuArgTyrrAlaThrAsnGlyAlaGlyAsnThrIle 628
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1705 -----CTTGGCAACCTTAAGAGCTTAGAGCT 1730
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 628 uProAsnIleSerLeu-----ThrIleProGlyValIleGlyLysProPro 643
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1731 ACTAATATCTTCCCTTAACATCTTATCGGCGCAATCCAGCAAGTATCGGGAGATCTACC 1790
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 643 C-----GlnArgLeuAsnAsnThrPheSerGlyThrAsnTyrrAsnAsnLeuGlnTyrrGlu 661
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1791 TCTACTACAGCAAGCTGAGAC-----CTTCTTAACAATTAATCTCAAA---GG 1832
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 661 YAsp-----PheGlyTyrrPheGlnPheProSerThrValIThrLeuProLeuAsnArg 678
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1833 AGAATATACCAAGATTAATTAATTCAGAAATGCTCATCTGTTATCTTGAAGGCAATAG 1892
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 678 gAsnIleProPheIlePheAsnArgAlaAspValSerAsnSerIleLeuIleIleAspLys 698
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

Db 1893 GGAGCTATG-----CGGAGAGTATGATCTCCATATGCAATC----- 1931  
 QY 698 atleaguphelleProlethSerSerethesglinaamargluylugluylleucl 718  
 Db 1932 -----GTGCCCCCAAGTTTCTCAAGATAGAAAGAAAGAAATTGA 1975  
 QY 718 urthileglnthlyleasnthrPhepetherAenhiethr 732  
 Db 1976 CCAGATTGTTGATCCATATAGTTGGTTTCTTGCTACTCACTAC 2018

RESULT 7  
 AY416726 1986 bp DNA linear GSS 17-DEC-2003  
 LOCUS AY416726  
 DEFINITION Mus musculus LY64 gene, VIRUTAL TRANSCRIPT, partial sequence,  
 genomic survey sequence.  
 ACCESSION AY416726  
 VERSION AY416726.1 GI:39772686  
 KEYWORDS GSS.  
 SOURCE Mus musculus (house mouse)  
 ORGANISM Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;  
 Sciurognathi; Murioidea; Muridae; Murinae; Mus.  
 1 (bases 1 to 1986)  
 Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarimal,A.,  
 Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,  
 Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Shtinsky,J.J.,  
 Adams,M.D. and Cargill,M.  
 Inferring nonneutral evolution from human-chimp-mouse orthologous  
 gene trios  
 JOURNAL Science 302 (5652), 1960-1963 (2003)  
 PUBMED 14671302  
 REFERENCE 2 (bases 1 to 1986)  
 Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarimal,A.,  
 Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,  
 Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Shtinsky,J.J.,  
 Adams,M.D. and Cargill,M.  
 Direct Submision  
 Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,  
 Rockville, MD 20850, USA  
 COMMENT This sequence was made by sequencing genomic exons and ordering  
 them based on alignment.  
 FEATURES  
 source Location/Qualifiers  
 1..1986  
 /organism="Mus musculus"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:10090"  
 <1..>1986  
 /gene="LY64"  
 /locus\_tag="HCM5986"

ORIGIN  
 Alignment Scores:  
 Pred. No.: 0.143 Length: 1986  
 Score: 117.50 Matches: 114  
 Percent Similarity: 31.7% Conservative: 76  
 Best Local Similarity: 19.0% Mismatches: 190  
 Query Match: 3.0% Indels: 221  
 DB: 10 Gaps: 22

US-10-783-417-2 (1-735) x AY416726 (1-1986)

QY 201 AenValhAaBnaBpHeileargluileProgl----- 212  
 Db 112 AACTGTGAATAATTAGCTCAATGAATTCCTGGACACTTACCAACTCAACAGATGT 171  
 QY 213 -----PheglneuglnthrylyrthlrleuleuleuProiletyrthlagnlaaia 230  
 Db 172 TTGAGCTTCAAGCTTTATGTCTTGCTCACTTCAAAACAGACCTTCAGCAGACTTATC 231  
 QY 231 AenPhehileuBnleuleuInglnlyAlagluleuAlaBpGlu----- 246

Db 232 AATCT-CACCTTCTGTGATTTAAACGAGTGCCAGATTTACTGATATGAAATACTTT 290  
 QY 246 ----- 246  
 Db 291 CCAAGCAACATCGGTTAGACACCTTGTTACTAACCGCAATTCCTGTATTTATGCG 350  
 QY 247 -----TTPAaAlaBpIlehiProSerglnilegluProaAlagly 261  
 Db 351 AAGACAGACCTTATGGGCTTAAGGCATTAAGCATGTGTTCTTCATCAAAACGAAT 410  
 QY 262 -ThSerBapPtyrTylyleuleulegluBnillePolyrTySerBnTyCy 281  
 Db 411 ATCCAGTATTTATTCCTCCACTGCACATCAAAAAACCTTGAA--AGTCTCATCT 467  
 QY 281 sAlaBnThryrAagThnglyleuBnleuAargBpGluProaBnMetlyerPse 301  
 Db 468 CGAAGCAACATATTTCTTCATTAAGTCTCCCAAGGTTTCCCAACAGAACTGAA 527  
 QY 301 rIlePhe-----AenBpTyraArgtyrMetThrIleThrValIleuBpThrl 318  
 Db 528 GGTCTCGATTTTCAGATATATGCTATTCATTCCTGTCTTAAGAAATATGAGCTCTCT 587  
 QY 318 eSerglnPheSerleuTyraBpIleuAargtyrAargBpSerileglylylelu-- 337  
 Db 588 CCAGCAAGCCACTAATCTGAGCTTAACCTTAATGAATGACATTCAGCAATAGAGCC 647  
 QY 337 ----- 337  
 Db 648 GGGGGCTTTCAGCTGCTGTCTCCAAAGTTAACTTGGAGGACTCAGAACTTGCT 707  
 QY 338 -----VallysglylelyeAaBgluBnleuThraargluileTyThrThrgluileA 355  
 Db 708 AGTATCTTCAAGGTTTGAAGAATCTTACATCCAGTCTCTTGCTGGGAGCAATTGA 767  
 QY 355 n-----PheaBpArgleuProglneulya 364  
 Db 768 GCACATGATGACGAAGATATTAGTCTGCGGTGTGAGGCTCTGTGAATAGTCTGT 827  
 QY 364 lclnProaBnleuAlaThrMetGluTy-----AenleuThraAAlaSerPheyle 382  
 Db 828 GGAGAGCATCAACCTAAGACATTAATTTCTTCAACATTTCTCCACACATTCATTC 887  
 QY 382 uPheSerPheleuglnPhe---IlePheTyThrgluAaBnThraBpGlyAaBn 401  
 Db 888 CTTCAGTGGCTTCCAGAACTGACCTTAACACCTACCTCAGTGAATGGCTTCTGG 947  
 QY 401 gHeuValglyleSerAaBnAargBpAlaProThryrSerAaBnThrlleThrgluThrle 421  
 Db 948 ACTGTGGAGCTAAGC----- 963  
 QY 421 uTyrglyluAargThnglySerProThrThrylThrlleargProPheGlu-SerTyrl 441  
 Db 964 -----ACACTTAAGAAATTAAGTTCTCAGTGC 989  
 QY 441 yValSerle---ValThraBpArglnSerProProvalSerProIleGlnProhIa 460  
 Db 990 AATATGATTGGAATTTCTGCAATATGACGTCTTCCAAATTCCTCCCTC-CTTACTCAC 1048  
 QY 460 heileleBnGlnilegluBnleuTyrlleuAaBnglySerSerAaBnThrlleuTylys 480  
 Db 1049 TTTCATC-----AAGGCAACAGAAAGAGACTTG 1078  
 QY 480 eAlaGly---GlySerleuSerAaBnTyrglnAaBnThrThrPheBpGlnPhePro---- 497  
 Db 1079 AGCTGTGATCTGCTGTGTAAGAACTTGAAGAAATCTCCGCAACTTGAACCTCAGCCATG 1138  
 QY 498 -----ArglylyAaBpCyAaBnleuValIleAaBpProglTySerProAaBnPha 515  
 Db 1139 ATGACATTAACATCTCTATGT-----TGCACCTTAACATC 1177  
 QY 515 aBnAaBnTySerhileuSer-----HisPheSerleuPhe 528  
 Db 1178 GAACCTGTCTCACTTAAGAGCTGAACTTGAAGTAAATGAAACCCCTGAGCCTCAAGA 1237

QY 528 hrTyrSerTyrValIleGlyLeuGlnIleLeuAspThrGly----- 543  
 Db 1238 CTGAGGCAATCAAGAATGCTCCTCAGCTAGACTCTGATTTGGCATTACTGCATTA 1297  
 QY 544 -----ValleuGlyTPT 548  
 Db 1298 AGGTGAAGATGACAGAGTCCCTTCAGAACCTTCATTTTGAAGGCTGAATCTCT 1357  
 QY 548 hrHisSerSerValAspArgTyrAsnAlaIleSerAspIleIleThrMetIleProA 568  
 Db 1358 CCCACAGCTCCTTATC-----ATCAGCAGTGAAGAGCTCTTCATAGGCTGCCAG 1408  
 QY 568 IAlle-----LysGlyAsnAsnLeuAspThrA 577  
 Db 1409 CACTCCAGCATTTGAACCTTACAGAGGAATCATTTCACAAAGGAATATCCAAAGACCA 1468  
 QY 577 snSerLysValIleGluGlyProGlyHisThrGlyGlyAsnLeuValTyrLeuGlnSerG 597  
 Db 1469 ACTCA-----CTTCAGACAC 1483  
 QY 597 lngIArgLeuGluIleThr-----CysGluThrProAsnSerThrGlnSerT 613  
 Db 1484 TGGGAAGACTAAGAAATCTGCTTTTATCTCTTTGATCTCTCTCCATTCACACGACAG 1543  
 QY 613 YrPheIleArgLeuArgTyrAlaIthr----- 621  
 Db 1544 CTTTCACACGCTGAAAGATGATGATCATGTAGACTGACACAGGCTGACATCCA 1603  
 QY 622 -----AsnGlyAlaGlyAsnThrL 628  
 Db 1604 GTAGCATTTGAGGCTCTTACTGATCTTAAAGGAGATCTTCACTCAATCTGGCTCAAT 1659  
 QY 628 euProAsnIleSerLeuThrIleProGlyValIleGlyIleProProGlnArg 645  
 Db 1660 -----CACATCAGCATCTCTTACCCAGTCTCTCCCATCTTGTCCACGACAG 1707

RESULT 8 2157 bp mRNA linear HTC 03-APR-2004  
 AK089255  
 LOCUS AK089255  
 DEFINITION Mus musculus NOD-derived CD1c +ve dendritic cells cDNA, RIKEN full-length enriched library, clone: F630107B15 product: lymphocyte antigen 78, full insert sequence.  
 ACCESSION AK089255  
 VERSION AK089255.1 GI:26354375  
 KEYWORDS HTC; CAP trapper.  
 SOURCE Mus musculus (house mouse)  
 ORGANISM Mus musculus  
 BUKERYOTA; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE  
 AUTHORS Carninci, P. and Hayashizaki, Y.  
 TITLE High-efficiency full-length cDNA cloning  
 JOURNAL Meth. Enzymol. 303, 19-44 (1999)  
 PUBMED 10349636  
 REFERENCE  
 AUTHORS Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Kono, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.  
 TITLE Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes  
 JOURNAL Genome Res. 10 (10), 1617-1630 (2000)  
 PUBMED 11042159  
 REFERENCE  
 AUTHORS Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Kono, H., Akiyama, J., Nishi, K., Katsunai, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujitake, S., Inoue, K., Ozawa, Y., Izawa, M., Ohara, B., Watabiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.  
 TITLE RIKEN integrated sequence analysis (RISA) system-384-format sequencing pipeline with 384 multicapillary sequencer

JOURNAL Genome Res. 10 (11), 1757-1771 (2000)  
 PUBMED 11076861  
 REFERENCE  
 AUTHORS 4  
 TITLE The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium.  
 JOURNAL Functional annotation of a full-length mouse cDNA collection  
 REFERENCE  
 AUTHORS Nature 409, 685-690 (2001)  
 TITLE 5  
 JOURNAL The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.  
 TITLE Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs  
 JOURNAL Nature 420, 563-573 (2002)  
 REFERENCE  
 AUTHORS 6 (bases 1 to 2157)  
 Adachi, J., Aizawa, K., Akimura, T., Arikawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, M., Hayashida, K., Hayatsu, N., Hiramoto, K., Hirooka, T., Hirozane, T., Hori, P., Imorani, K., Ishii, Y., Itoh, M., Kagawa, I., Kanakawa, T., Katoh, H., Kawai, J., Kojima, Y., Kondo, S., Kono, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numasaki, R., Ohno, M., Ohsato, N., Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shingawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, P., Takaku, Akahira, S., Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y.  
 TITLE Direct Submission  
 JOURNAL Submitted (16-APR-2002) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suenho-cho, Tsunumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-ree@gscc.riken.jp, URL: http://genome.gsc.riken.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)  
 COMMENT cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.  
 Tissues were provided by Dr. John Todd (Dept. of Medical Genetics Wellcome Trust Centre for Molecular Mechanisms in Disease Wellcome Trust/MRC building Addenbrookes Hospital Cambridge) whose assistance we gratefully acknowledge.  
 Please visit our web site for further details.  
 URL: http://genome.gsc.riken.jp/  
 URL: http://fantom.gsc.riken.jp/.  
 FEATURES  
 source location/Qualifiers  
 1..2157  
 /organism="Mus musculus"  
 /mol\_type="mRNA"  
 /strain="NOD"  
 /db\_xref="PANTOM:DB:F630107B15"  
 /db\_xref="taxon:10090"  
 /clone="F630107B15"  
 /cell\_type="NOD-derived CD1c +ve dendritic cells"  
 /clone\_idb="RIKEN full-length enriched mouse cDNA library"  
 134..2119  
 /note="unnamed protein product: lymphocyte antigen 78 (MGI:1194924, GBLNM\_008533, evidence: BLASTN, 99%, match=2136)  
 putative"  
 /codon\_start=1  
 /protein\_id="BAC40816.1"  
 /db\_xref="GI:26354376"  
 /translation="MAADPISCPVVALPASCRAATSSDOKIRKYNKYNCNIGLNEIPGLNNTSCBESFPNVLPPTONTTFSLININTFDLRQCIYVHBTPOGQR LDTVLVTNPLIPAEATLSCPKALKHLFPIQTGISIDFIPLNOKTLSLYASNH ISSILPKGPFTEKLVLDQPNNAIHYISKDMSIQAATISLNTKNDIAGIELGA FSNVAFQSNFGQNLIVIFKGLSGITGLMTGTFEDMDQDSDIPVFGQLCMVS ESNIOKHYFNISNTNFCPSGQGLDITLTHSGLSPSGVGUSTLKVLVSAKFS NLCOISANFSLTHLSIKNTKRLLELGTCLNENENREHLDSDIDETDCCNLOL RNHSLDLSNLSYVPSLSKTFARFCQQLDLDAFLRYKDAQSPQNDHLKVL NLSHSLDISSEQLFDGLPALQHLNLSQNHPPKNIQKNSLQTLRLLELIVLSPCDL







OY		IleGluProAsnAlaGlyThrSerAspArgTyrTyrLysLeuLeuLysGlnAsnIlePro	275
Dd		TTGAATTTGGCAGCGAGAACAAGAACTGGCTGTCCGACCAAGCATGACGT	687
OY		LysTyrSerAsnTyrCysAlaAsnThrTyrArgThrGlyLeuLysAsnLeuArgAspGlu	295
Dd		AGGATTCTTAATTCGTCGACGAGATAGTTGGCAACAGATATGCATTCGATGATAC	747
OY		ProAsnMetLysTrpSerIle---PheAsnAspTyrArgArgTyrMetThrIleThrVal	314
Dd		AACCTTGAGAGATACCCTGGTTCGAAATTCACAGAAATATCTAAGATCATGCACCTTCGTCGA	807
OY		LeuAspThrIleSerGlnPheSerLeuTyrAspIle-----	326
Dd		CTTCTCAGCTGATGAAGATTTGTGTTGCAGAAAATAGCTTTCTGTGTGCCAATAA	867
OY		-----	326
Dd		ACTACTGAATCTTGCTACATCATAGTCTCCCCCTACTAATGTGGGTCTCATGAT	927
OY		LysArgTyrArgAspSerIleGlyGlyIleGluValLysGlyIle----LysAsnGlu	344
Dd		AAAGAGTTAGAAAGTTTCATGGGCCCAAGTGTATACGAAATATCATCTTTCACATAG	987
OY		LeuThrArgGluIleTyrThrThrGluIleAsnPheAspArgLeuProGlnLeuArgVal	364
Dd		CTGAGAGAGGCTGTAAAAAGCTCATCATTTGGAAATGTGCACATGTGATCTTCAAA	1047
OY		GlnProAsn-----LeuAlaThrMetGluTyr	373
Dd		TGTCCAATGCACATTTAGCATGTCGCAATGACTGTGTGAAGTTGGCATTTCCAGAAATG	1107
OY		AsnLeuThrArgAlaSerPheLysLeuPheSerPheLeuGlnGlnPheIlePheTyrThr	393
Dd		AGGATATTACAGAGCTCTTTGTGTTGCAATTTGACATTTGCAGAACCTGTATATCTGCTA	1167
OY		GluAsnThrAsnPheGlyAsnArgLeu---ValGlyIleSerAsnArgAspAlaProThr	412
Dd		CTGACTTGTCATGGATTAATGAATGATTTGCTATTTGATTCAGACATTAAGCACAA	1227
OY		TyrSerAsnThrIleIleThrGluThrLeuTyrGlyGluArgThrGlySerProThrThrLys	432
Dd		ATGCATTACGTTATCTTCCAGATTTGAAGGACG-----GGGCGAGTGAAGCCAAAG	1281
OY		ThrIleArgProPheGlnSerTyrLysValSerIleValThrAspArgGlnSerPro	452
Dd		CTCACAGATATTA-----GCACAGATCCGCTA	1311
OY		ValSerProIleGlnProIlePheIleIleAsnGlnIleGluLeuTyrLeuAsnGlySer	472
Dd		TGTGATACCAATGAACTCATATATCTCAAAATATATCAAATGATATTTGGAAGGCGAGA	1371
OY		SerAsnAsnThr-----LeuLysTyrSerAlaGlyLysSerLeuSerAsnTyr	488
Dd		AATGATATTCGATGTGACCTGTGAAAGATTCATTCGAAAGACCTCTTTGTTAT	1431
OY		GlnAsnThrThrPhePheGlnPheProArgLysLysAspCysAsnLeuValIleAspPro	508
Dd		ATAAATGCACCTTATTCACAGATAT	1467
OY		GlyCysSerProAsnPheAsnAsnTyrSerHisIleLeuSerHisPheSerLeuPheThr	528
Dd		GGTATT-----GGTTATCGCGCTATGCTCAGCATTAAGAGTAATGTAATATA-----	1515
OY		TyrSerTyrValIleGlyLeuGlnLeuGlnIleLeuAspThrGlyValLeuGlyTyrThr	548
Dd		ACAACTTCCTATGTTTAAGGAGAGCAACCTTACCATGGCGCTCTTGCAACTTGT	1575
OY		HisSerSerValAspArgTyrAsnAlaIleSerAspLysIleIleThrMetIle-----	566
Dd		GGCTCTTCACTTAG-----TTCTTGATATTTTCCAGACCGCGTTCATGATATTA	1632
OY		---ProAlaIleLysGlyAsnAsnLeuAspThrAsenTyrValIleGlyLysProGly	585

Db	1633	GCACCACTTATTAAGAAGAAAT-----TCCTACTTGAAGATGTCTGAAAGCAGCTGGA	1683
Qy	586	-----HisThrGlyGlyAsnLeuVal	592
Db	1684	TGCCGTAACTTGCTATTTGAAACATGCTGAAGTAGAAGCAGATGAGTGTGGTAAATATATATAC	1743
Qy	593	-----TyrLeuGlnSerGlnGlyArgLeuGlnIleThrCysGlnThrProAsnSerThr	610
Db	1744	GGAGATATTTTCTACAA-----GAAATTAACAGTACTTGCTGCTAGAGAGATGTGAA	1794
Qy	611	GlnSerThr-----PheIleArgLeu	617
Db	1795	ATGGAGATGGCAGTTTGGCCCTATTCCGATTACCAACCTTATCTCTTTTAGTAAGTA	1854
Qy	618	ArgTyrAlaThrAsnGlyAlaGlyAsnThrLeuPro---AsnIleSerLeuThrIlePro	636
Db	1855	AGAAAGATGACTATATGGACTTGTCGACAAACATTACAGAGATATATATGTGNGCTTCTCT	1914
Qy	637	GlyVal	638
Db	1915	GATATT	1920
RESULT 11			
CNS09YP2	LOCUS	CNS09YP2	2460 bp mRNA linear HTC 06-FEB-2004
DEFINITION	Arabidopsis thaliana Full-length cDNA Complete sequence from clone GSJLTPGH72A12 of Hormone Treated Callus of strain col-0 of Arabidopsis thaliana (thale cress).		
ACCESSION	BX832541 GI:42455129		
VERSION	1		
KEYWORDS	HTC; GSJLT cDNA.		
SOURCE	Arabidopsis thaliana (thale cress)		
ORGANISM	Arabidopsis thaliana		
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsids.		
AUTHORS	1 (bases 1 to 2460)		
REFERENCE	Castelli,V., Aury,J.M., Jallion,O., Wincker,P., Clepet,C., Menard,M., Cruaud,C., Quetier,F., Scarpelli,C., Schachter,V., Temple,G., Caboche,M., Weissenbach,J. and Salanoubat,M.		
TITLE	Whole Genome Sequence Comparisons and 'Full-length' cDNA Sequences: A Combined Approach to Evaluate and Improve Arabidopsis Genome Annotation		
JOURNAL	Unpublished		
REFERENCE	2 (bases 1 to 2460)		
AUTHORS	Genoscope.		
JOURNAL	Direct Submission		
COMMENT	Submitted (18-NOV-2003) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr)		
	The sequences are based on single pass reads.		
	Life Technologies (a division of Invitrogen) members carried out full-length libraries construction : Temple G.		
	Genoscope members carried out sequencing and annotation : Castelli V., Aury J.M., Jallion O., Wincker P., Menard M., Cruaud C., Schachter V., Weissenbach J., Salanoubat M.		
	URGV INRA : Clepet C., Caboche M.		
	Annotation is based on the June 2003 version of the Arabidopsis genome released by MIPS (Munich Information center for Protein Sequences). 5 prime and 3 prime are assembled with Phrap.		
	http://www.genoscope.cns.fr/externe/sequences/Banque_Projet_EF/Full length		
FEATURES	http://www.genoscope.cns.fr/cgi-bin/ggb/ggb?source=Arabidopsis.		
SOURCE	Location/Qualifiers		
	1..2460		
	/organism="Arabidopsis thaliana"		
	/mol_type="mRNA"		
	/db_xref="taxon:3702"		
	/clone="GSJLTPGH72A12"		
	/tissue_type="Hormone Treated Callus"		
	/ecotype="Col-0"		
	/plasmid="pCMVSPORT_6"		

gene 1..2460  
/gene="A15g20250"

ORIGIN

Alignment Scores:

Pred. No.:	0.397	Length:	2460
Percent Similarity:	115.00	Matches:	170
Best Local Similarity:	31.7%	Conservative:	109
Query Match:	19.3%	Mismatches:	269
	3.0%	Indels:	332
		Gaps:	49

US-10-783-417-2 (1-735) x CNS091P2 (1-2460)

```

QY 57 ThrGlnTYGlyAspAsnPhleglu-----ThrPheAlaSerAlaAspThrIle 72
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 116 ACGAGGTACAGAGAAAACGAGAGAGAGAGATATGACATTAACCGCG----- 166
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 73 AlaAlaValSerAlaGlyThrIleVal-----SerGlyThrLeuAlaGlyIleGly 90
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 167 GTTCGATATTCGACGCGGAATCTAATCATCAAAAACCGACGATCTAACCGGTGACCA 226
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 91 Gly-----LeuThrSerIleSer-----GlyProIle---GlyIle---IleGly 103
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 227 GATTAATGTCATACGACGTCAGCATCGAGACTGACCGGTAGAAAGAGACTTTCGCGCA 286
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 104 AlaIle-----IleIleSerPheGlyThrLeuIleThrVal 115
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 287 GCTGATATTTAACAAGAGAGAGTAACACATCGTACCGATCGGTACGCTTCGCAATTC 346
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 116 PheTTPProAlaGlyGlnGlnAspIlyThrValTTPThrGlnPheIleIysMetGlyIle 135
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 347 CCGTTATGTCCTGTTCCGCTTTAAG---CTCGGTGATGAGCCGACAGATGAGAGAA 403
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 136 IlePheValAspThrPro-----LeuThrGluSer----- 145
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 404 ATGGGCCGAGATATTCCTTACAGACGCACTTATTTGTGTCAGAGCAACATGGGTTC 463
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 146 -----IleIysGlnLeuIys 150
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 464 CACCTTGAGTCTGACGAGCTTAACGCGTCGAGCTGAACCAAGATTACACCGTTTC 523
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 151 LeuGlnThrLeuGlnGly---PheArgGlnIleLeuGln----- 162
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 524 TTGCCGTTATCGAAGATCTTCCGTTGCTGTCACAGAAACGTTAACGATGAGATT 583
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 163 -----SerTyrAsnThrAlaLeuAspAspTirArgIlyLeuIys 175
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 584 GAGCTCTGTTGGAAGTGTGACGTCGACACTTAAACGTCGCTTACTCACTCTCTG 643
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 176 ArgLeuGlnAlaProGlyLeuProProSerSerAlaLeuGlnGlnAlaIleLeuThrLeu 195
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 644 TATATTCACGCC---GGTACAGATCCGTTTCAGACATACGAGCGCTATTGGCACCTT 700
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 196 IysIleArgPheGlnAsnValHisAsnAspPheIleArgIleIleProGly----- 212
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 701 AAGTTCATCTGAATAGTTCCGTCACACGTCACGAAAGAGACTTCCAGAGATCGTTGAC 760
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 213 ---PheGlnLeuGlnThrTyrIlyThrLeuLeuLeuProIleTyrAlaGlnAlaIleAsn 231
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 761 TACTTCGATGTGTGACCTTGGAACGCGTTTATCAAGAGTGACTCAGAAAGCGCTGAA 820
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 232 PheIleLeuAsnLeuGlnGlnGlyAlaGlu-----LeuAlaAsp 245
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 821 GGTGATCTTAATGCTCTCCGCCGCGGTGTACGCCGCCGAAGTTGTTATCATACACAC 880
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 246 GluTTPAsnAlaAspIleHisProSerGlnIleGlu----- 257
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 881 GGTGGCAATCAGTTGAGAGAGATGACACGCTGAGGCGGAGATGAGAGAAAGAGTCA 940
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 258 -----Pro 258
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 941 CCGATTTTTCAGACTGACGGGATCAAGAGAAAGAAAGTTTAAAGAGAGATGATCCA 1000

```

```

QY 259 AsnAlaGlyThrSerAspAspTyrTyrIlySerLeuLeuGlu----- 272
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1001 AACGTTGGGATTAAG---AACATGTGCAAGATGTGTAAGAGAGAAACACGTTTGAATAT 1057
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 272 ----- 272
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1058 GTTTACGTGTGACACGCAATACCGGTTATTGGGCGGAGTTTACGCCGGAGAGATAC 1117
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 273 ---AsnIleProIlyTyrSerAsnTyrCysAlaAsn-----ThrTyr 285
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1118 GATCTGTATGAAATATCTTATATGTCAAAAGGTGTGTGAGAGATGATCCAGCTGG 1177
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 286 ArgThrGlyLeuIysAsnLeuArgAsp-----GluProAsnMetIlySerSer 301
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1178 AAGACTGATGATTAATACCGCTGCAAGGTTGGGTTGTGTTAGCCCAAGAAAGTATATAG 1237
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 302 IlePheAsnAspTyrArgArgTyrMetThrIleThrValLeuAspThrIleSerGlnPhe 321
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1238 TTTTACAATGAGCTTCATAGTACTTGTGCTGATGCTGGCGTGGACGCTGTGAAA----- 1291
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 322 SerLeuTyrAspIleIysArgTyrArgAspSerIleGlyGlyIleGluValIlyGlyIle 341
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1292 -----GTGATGTCAAGTGTATTTGAGACTTGGGTGT-----GTTTACGCGGT 1339
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 342 LysAsnGlnLeuThrArgGlnIleTyr-----ThrThrGlnIleAsnPhe 356
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1340 CCGGTTGAGCTGACCTCGTACAGTTTCATCAAGCTCTTGACTCCTCGTTGCTAAGAACTTC 1399
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 357 AspArgLeuProGlnLeuArgValGlnProAsn-----LeuAlaThrMetGluTyr 373
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1400 -----CCGATATATGCGTCGATCTGTATGGAGCCAT 1432
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 374 Asn-----LeuThrArgAlaSerPheIlySerLeu 382
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1433 AATACGATGCTCTTACTGCTCGAAGCAAGAGCTGTGATTAAGAGATCAATGAT--- 1489
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 383 PheSerPheLeuGlnGlnPheIlePheTyrThrGlnAsnThrAsnPheGlyAsnArgLeu 402
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1490 -----TTCTATCA----- 1498
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 403 ValGlyIleSerAsnArgAspAlaProThrTyrSerAsnThrIleThrGluThrLeuTyr 422
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1499 -----CGGATCCGGTGTCTCATACCATCCATATGACCTCCGTTCTTAT 1543
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 423 -----GlyGlu-----ArgThrGlySerProThrThrIlyThrIleArg 435
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1544 AACAGTGTCTTGGAGAGATTATGACAGCTGACTGGAGATATGTTCCATTCCGTGCAC 1603
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 436 PropheGluSerTyrIlyValSerIleValIThrAspArgGlnSerProProValSerPro 455
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1604 CCTGTGCAAGATATCATGACCTCTCTGAGGCCATCATGCGGGGCT----- 1651
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 456 IleGlnProHisPheIleIleAsnGlnIleGluLeuTyrLeuAsnGlySerSerAsnAsn 475
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1652 -----CTATATGTCAGAGATTCTCTGGA--- 1675
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 476 ThrLeuIyTyrSerAlaGlyIlySerLeuSerAsnTyrGlnAsnThrThrPheGln 495
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1676 ---AAGACAAAC-----TTTGAG 1690
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 496 PheProArgIly--- 499
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1691 CTTTCAAGAAAGCTGTATGCTGTATGATGATGATTTCTGTCGTCGACTCCCTGTAGA 1750
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 500 -----LysAspCysAsnLeuValIleAspPro-----GlyCysSerPro----- 512
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1751 CCAACTCCCGCATGTG---TTGTTCCGTGATCTCTGCGCCGTGATGCTGACGCTTGTAAAG 1807
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 513 -----AsnPheAsnAsnTyrSerHisIleLeuSerHisPheSerLeuPheThrTyrSer 530
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1808 ATATGAAACATGAAACAAGTAC----- 1828

```

FEATURES	Location/Qualifiers
QY	531 TyrVal11IeGlyLeuGlnLeuGln11LeuAAspThrGlyVal11LeuGly----- 546
Db	1829 -----ACTGGAAGTCTTGCGGTGATTAAGTCC 1855
QY	547 -----TrpThrHisSerSerValAspArgTyrAsnAlaIle---SerAspLys 561
Db	1856 CAGGAGGACGAGCTGG-----AGCGACACAGAGAGAAAAACATTTCACACAGACTTAA 1909
QY	562 IleIleThrMetIleProAlaIleLeuGlyGlyAsnAsnLeuAAspThrAsnSerLysValIle 581
Db	1910 ACTGATACCTTCATCGCTCCATTCGTGGTGGTATGTGCATTCAATATCGAAGGCTCC 1969
QY	582 GlnGlyProGlnHisThrGlyGlyAsn---LeuValTyrLeuGlnSerGlnGlyArgLeu 600
Db	1970 ACTGATGACCAACAACCTGGAATGGAATGAACTGCTGCTTTTAACTCCAGAGAGAGGCGAACTT 2029
QY	601 GlnIleThrCysGluThrProAsnSerThrGlnSerTyrPheIleArgLeuArgTyrAla 620
Db	2030 ATTGTGTAAGCCATACATACATGTGTCTTCCAGCTTCATCTCAAAATCCGAGACAGATC 2089
QY	621 ThrAsnGlyAlaGlyAsnThrLeuProAsnIleSerLeuThrIleProGlyValI1IeGly 640
Db	2090 TTC-----ACGGTAGGCCCATTAAGTCAATCTGTGGATGGTATCTTTT 2134
QY	641 IleProProGlnArgLeuAsnAsnThrPheSerGlyTyrAsnTyrAsnAsnLeuGlnTyr 660
Db	2135 GCCCAATATTGCTGTAGTAAACATGTACATTCGGAGAGAGCTATCGAAGAGCTTAGAAT 2194
QY	661 -----GlyAspPheGlyTyrPhe 666
Db	2195 GAAGCCGAGAAAGATGAAGAAGTGTAAAGATGAAGATGAGATGGCAATTCGATCTTAC 2254
QY	667 GlnPhe-----ProSerThrValThrLeuProLeuAsnArgAsnIleProPheIlePhe 684
Db	2255 TCTTCTGTGAAGCCTTAAGATGCGTGTGTGAGCAAT---GAGATTGCATTCGAGATAC 2311
QY	685 AsnArgAlaAspValSerAsnSerIleLeuIleIleAspLysIleGluPheIleProIle 704
Db	2312 GAT-----TCTCTCTGTGATGTGTCACCTTGTGAGTTAGACAAATATGCTTATA 2359
QY	705 ThrSerMetHisGlnAsnArgGlnLysGlnLysLeuGluThrIleGlnThrLysIle 724
Db	2360 -----GAGAACAAACGATTTTCATCTGATCCAAAGTTGAGTTA 2395
RESULT 12	
LOCUS	CFP951415 820 bp mRNA linear EST 20-NOV-2003
DEFINITION	UI-M-HL0-cnb-o-16-0-UI_r1 NIH_BMAP_HL0 Mus musculus cDNA clone IMAGE:30633855 5', mRNA sequence.
ACCESSION	CFP951415
VERSION	CFP951415.1 GI:38467284
KEYWORDS	EST.
SOURCE	Mus musculus (house mouse)
ORGANISM	Mus musculus Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muridea; Muridae; Murinae; Mus.
REFERENCE	NIH-MGC <a href="http://mgc.nci.nih.gov/">http://mgc.nci.nih.gov/</a> .
AUTHORS	National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE	Unpublished (1999)
JOURNAL	Contact: Robert Strausberg, Ph.D.
COMMENT	Email: <a href="mailto:cgabbs-remail.nih.gov">cgabbs-remail.nih.gov</a> Tissue Procurement: Dr. James Lin University of Iowa cDNA Library preparation: Dr. M. Bento Soares, University of Iowa cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa DNA Sequencing by: Dr. M. Bento Soares, University of Iowa Clone Distribution: Distribution information can be found at <a href="http://genome.wiowa.edu/distribution/mouseefl.html">http://genome.wiowa.edu/distribution/mouseefl.html</a> This clone was contributed by the Brain Molecular Anatomy Project (BMAP)
Seq primer:	pYX-5.

[illegible]



Db 226 ACCCCCTTTTGGGAAAAAACACTCCCTCCCTTCACTA-AACCTACTTCTT 168  
Qy 681 -----ProPhelIephEaAnrAlaAspValSerAsnSerIleu 695  
Db 167 AATAAAACCCGCCACCTTTTGG-----AAAAATACCTCTCC 123  
Qy 695 IeIeAspLyIleGluPheIleProIleThrSerMet 708  
Db 122 GCCCTGGCCCTTTTATTTATTTACACCCATTACTCTCTCC 82

RESULT 14  
CL974457 2841 bp DNA linear GSS 21-SEP-2004  
LOCUS CL974457  
DEFINITION O1PCC025896 Oryza sativa Express Library Oryza sativa (indica cultivar-group) genomic, genomic survey sequence.  
ACCESSION CL974457 GI:52403441  
VERSION CL974457.1  
KEYWORDS GSS.  
SOURCE Oryza sativa (indica cultivar-group)  
ORGANISM Oryza sativa (indica cultivar-group)  
Bukayota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Eriarthodeae; Oryzaceae; Oryza.  
REFERENCE 1 (bases 1 to 2841)  
Ma, L., Wang, J., Chen, C., Liu, X., Su, N., Li, L., Wang, X., Cao, M., Jiao, Y., Sun, N., Zhang, X., Bao, J., Sun, D., Zhao, H., Yuan, L., Wong, G. K. S., Deng, X. W. and Wang, J.  
An analysis of transcriptional regulation of the rice genome and its comparison to Arabidopsis  
Unpublished (2004)  
JOURNAL Department of Bioinformatic  
COMMENT Contact: Chen Chen  
Beijing Institute of Genomics  
Chinese Academy of Sciences, Beijing 101300, China  
Tel: 86-10-80481559  
Fax: 86-10-80488676  
Email: chenchen@genomics.org.cn  
Rice genomic sequence.  
Class: exon-trapped.  
Location/Qualifiers  
FEATURES  
source 1..2841  
/organism="Oryza sativa (indica cultivar-group)"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:39946"  
/clone\_lib="Oryza sativa Express Library"  
/note="Oryza sativa exon trapped genomic sequences"

ORIGIN  
Alignment Scores:  
Pred. No.: 0.982 Length: 2841  
Score: 112.50 Matches: 175  
Percent Similarity: 33.1% Conservative: 111  
Best Local Similarity: 20.2% Mismatches: 313  
Query Match: 2.9% Indels: 267  
DB: 10 Gaps: 47  
US-10-783-417-2 (1-735) x CL974457 (1-2841)

Qy 24 ArgAsnSerAsnAspSerArgTyPProTyThrAsnAsnProAsnGln-----ProLeu 41  
Db 499 CGCAATGAAGCCTTGCAAGAAATTTGCACATTAACCATATAGCGTAAATCCCTG 558  
Qy 42 GlnAsnThrAsnTyIleGluTrpLeuAsnMetCysGlnGlyAsnThrGlnTyIleAsp 61  
Db 559 GACTCTAC-----ATGCTTGAATATTAATTAGCAACGAGT 597  
Qy 62 AsnPhelIuThrPheAlaSerAlaAspThrIleAlaIleAlaValSerAlaGlyThrIleVal 81  
Db 598 AATTATGAAGCTAGAGGTATTAAGCAATATCAGCAAGTCT-----GGATGGGTATT 651  
Qy 82 SerGlyThrLeuLeuAlaGlyIleGlyGlyLeuThrSerIleSerGlyProIleGlyIle 101  
Db 652 GCC-----AAGAAATGTGAGGATAGCATATGACAGCCCATGACATTGGTTTC 699

Qy 102 Ile-----GlyAlaIleIleIleSerPheGlyThrLeuIleThrValPheTrpProAla 119  
Db 700 TTCTATCTGGAATGAGACCTTGGAATGAGAGCACTG----- 738  
Qy 120 GlyGluGlnAspLySerValTyPThrGlnPheIleLysMetGlyGluIlePheValAsp 139  
Db 739 -----AGCAATAGTATGATTTTGACT-----GAGCTTTTCTGAT 774  
Qy 140 ThrProLeuThrIleLeuSerIleLysGlnLeuLysLeuGlnThrLeuGluGlyPheAsnGln 159  
Db 775 AATTCAAGTGTGGCATCTTGAAA-----CTAACTTACAAGAAATGCAACCATATTGAGA 831  
Qy 160 IleLeuGlnSerTyIle-----AsnThrAla----- 167  
Db 832 TTATGCTTGGCTTACCTGCGCATCTTCCTTAAGGTACATATTAAGTAAAGCAAGCTG 891  
Qy 168 LeuAspAspTrpArgLySerLeuLysArgLeuGlnAlaProGlyLeuProProSerSerAla 187  
Db 892 GTTCACCAATGATGATGCTTGTTTATTTAG-----CATCTTAAACA 936  
Qy 188 LeuGlnGlnAlaAlaLeuThrLeuLysIleArgPheGluAsnValIleAsnAspPheIle 207  
Db 937 TTTTCATCCGTACGACTTGCGGAGAG-----TATATC 969  
Qy 208 ArgGluIleProGlyPheGlnLeuGluThrTyIleThrLeuLeuLeuProIleTyIleAla 227  
Db 970 AGCAACCTGTGGG-----ATGCTTTCCTTACAGCTTCCATATTTGACACTGAACA 1023  
Qy 228 GlnAlaAlaAsnPhelIleLeuAsnLeuLeuGlnGlnGlyAlaGluLeuAlaAspGluTrp 247  
Db 1024 GAAGTTTACCACTGACGATATGTTGATGATGTTGCGAGATCGGTCAATGATGAAG 1083  
Qy 248 AsnAlaAspIleHisProSerGlnIleGluProAsnAlaGlyThrSerAspAspTyIle 267  
Db 1084 TTGGTTTCTTTATATGACAAATAATGACAGTACACAGAACAAATAATTTGCCACTAT 1143  
Qy 268 LysLeuLeuLysGluAsnIleProLySerAsnTyIle----- 280  
Db 1144 GCTTTCCTT-----GAAATTAATGACAGTCAATCAATATATGCTGCGACG 1200  
Qy 281 -----CysAlaAsnThrTyIleArgThrGly----- 288  
Db 1201 TTAAGGCACTATACCTTGATTCGATTCAGCAAAATGTGTCTCAAGATGAAATTTC 1260  
Qy 288 ----- 288  
Db 1261 TTCACAAAGTCTTGCGTCTAGATTAACTGACTGCTCCATAAGATCTCCCAAGC 1320  
Qy 289 -----LeuLysAsnLeuArg-----AspGluProAsnMetLysTrpSerIle 302  
Db 1321 TCTATTGGCAAAATTTGAAGAGCTGAGTTCCTTATGCTCCAAACATGAGATTAATGTC 1380  
Qy 303 PheAsnAspTyIleArgTyIleMetThrIleThrValLeuAspThrIleSerGlnPheSer 322  
Db 1381 TTCCCA-----AAGATATTACTCTGCTCCCAAACTA----- 1413  
Qy 323 LeuTyIleAspIleLysArgTyIleArgAspSerIleGlyIleGluValLysGlyIleLys 342  
Db 1414 -----AATATCCTTGACCTTCATGATCTTTTGGAATTTCAAGATTGCA 1458  
Qy 343 AsnGluLeuThrArgGluIleTyIleThrGluIleAsnPheAspArgLeuProGlnLeu 362  
Db 1459 GGATCAATTAGCAAAATGCTGTCATATCATCTTGACCTATCTGTTGCTTAATATC 1518  
Qy 363 ArgVal-----GlnProAsnLeuAlaThrMetGluTyIleAsnLeuThrArgAlaSerPheLys 381  
Db 1519 AGAGTTATACAGCCAAAGCTATAT-----GGCTCACCAAACTGCAATTGTTG 1569  
Qy 382 LeuPheSerPheLeuGluGlnPheIlePheTyIleThrGluAsnThrAsnPheGlyAsnArg 401  
Db 1570 AACTTATCGTGGGCTCAATCTTCAATATAGCCAGAAATATTTCAAGTCTTACTGAA 1629

QY 402 Leu-----ValGlyIleSerAsn-----ArgAspAlaProThrTyrSerAsn 415  
 Db 1630 CGCATATATTAATCAATCAAACTGCTTCTTCAGCCAGTATCATCATATTTGG 1689  
 QY 416 ThrIleThgluThrleuTyrGlyGluGlyngly---SerProThrThrIleThrIle 434  
 Db 1690 ACCCTCACTGAACACTCAATATTTGAACCTATACAGGTTCAGGCGCTTTGTAATTTGCCA 1749  
 QY 435 ArgProPheGluSerTyrIleValSerIle----- 444  
 Db 1750 ARGTCCTTCGGAATCGAAGATTTGGTACATCTTGACCTGTCAGGCTGCTCCGGGTA 1809  
 QY 445 -----ValThrAspArgGlnSerProProValSerPro 455  
 Db 1810 CAAGATTTCAAAAGCTTTGGTGGCTCACCAAACTGACACTTAATCAATCAAG 1869  
 QY 456 Ile-----GlnProHisPheIle----- 461  
 Db 1870 ATATTGGTAGAACAGAGTGGGTGATTAATGGAGATGTTATCTGAACCTATCACTACT 1929  
 QY 462 IleAsnGlnIleGluIleuTyrLeuAsnGlySerSerAsnThrleuTyr----- 479  
 Db 1930 CTTAATGATCTTAG---TATCAAACTTCTCGGAATAGCAGATTTGATTTACCA 1986  
 QY 480 SerAlaGlyGlySerIleuSerIleuTyrGlnAsnThrPhePheGlnPheProArgIys 499  
 Db 1987 AGAAGCCTTGAACCTCAAAAAGCTTCAA---ACATGGACCTTCACTACTGCTGCT 2043  
 QY 500 -----LysAspCysAsnLeuValIleAspPro 508  
 Db 2044 CTGAGATCCCTCCACACAGCATAGATTAAATTGACAGTCTGAGTTCTAATAGTTGG 2103  
 QY 509 GlyCysSerProAsnPheAsnAsnTyr-----Ser 518  
 Db 2104 GGGTGTCTGATCACTTAAGAGATTTGAGAAATCTCATTCAGAAATATTCAGTCT 2163  
 QY 519 HisIleLeuSerHisPheSerIleuPheThrTyrSerTyrValIle-----Gly 534  
 Db 2164 CATTAAC-----CACACTCATGTTCACTATTTCTGCTCAAGAGACAGACAGGA 2217  
 QY 535 LeuGlnLeu-----GlnIleLeuAspThrGlyValLeuGlyTTPPhrHisSerSer 551  
 Db 2218 ATGAGCTTCAGAGAAACAAATCTTCAACATTCAGATTCATTTGAGCT---TCAAG 2274  
 QY 552 ValAspArgTyrAsnAlaIleSerAspIysIleIleThMetIlePro----- 567  
 Db 2275 GCTGACCGACTTTTGAGACAGAGATGTTTGGAGAACTAATGCCACCAAGAGCTCTG 2334  
 QY 568 -----AlaIleIleGlyAsnAsnLeuAspThrAsnSerIysValIleGluGlyPro 584  
 Db 2335 TGGATTATTAAGCATACAGGGTATGTATGACACAACTGCCAACTGTTTGAAGCATTC 2394  
 QY 585 GlyHisThrGlyGlyAsnLeuValTyrLeuGlnSerGlnGlyArgLeuGluIleThrCys 604  
 Db 2395 TCACATCATCTTCCAAAGCTCTTTAAGATTGAGTT- GAAGAGTTTGCACAGAG----- 2447  
 QY 605 GluThrProAsnSerThr----- 610  
 Db 2448 CAGACACCTGCCACCACTGGAGACAGTTCGCCAAACCTGAATGAATGCATCTCCACAGAT 2507  
 QY 611 ---GlnSerTyr-----PheIleArgLeuArgTyrAlaThrAsnGlyAla 624  
 Db 2508 GGAAGACTTCACAAATAATTGATGGCGTTTCTCGGTGGCAGCAA----- 2552  
 QY 625 GlyAsnThrleuProAsnIle---SerleuThrIleProGlyValIleGlyIleProPro 643  
 Db 2553 ---AGCGGATTCGCGAAACTGAAGTCTTCACTTGAAGAGAT----- 2594  
 QY 644 GlnArgLeuAsnAsnThrPheSerGlyThrAsnTyrAsnAsnLeuGln-TyrGly----- 661  
 Db 2595 -----GAAACAGTTGGAAGAGTGGAGCACACACACAGTAAGAGTGGTGA 2642  
 QY 662 -AspPheGlyTyrPheGlnPheProSerThrValThrleuProLeuAsn----- 677

Db 2643 GATATGATAGATGATCATGTTCTTCATGCTAGTACTGCTCAATAGAGCTTTGCCAAT 2702  
 QY 678 -----ArgAsnIleProPheIlePheAsnArgAlaAspValSerAsnSerIleIle 694  
 Db 2703 GCTTAGACTGAACACATGCGCCACCGATGTTTACAGCTGGCTCATATTCAGAGAGCA 2762  
 QY 694 uIleIleAsp-----LysIle-----GluPheIlePro 703  
 Db 2763 GGGTTCGACAGGGGTTTAACGAGAGGCTGTATGAAGATTTTCGCTGAAGATTTATGCC 2822  
 QY 703 oIleThrSerSer 707  
 Db 2823 TACCGACTCTTCA 2835  
 RESULT 15  
 CDS0A1J6  
 LOCUS  
 DEFINITION  
 Arabidopsis thaliana Full-length cDNA Complete sequence from clone GSLTFB30ZG08 of flowers and buds of strain col-0 of Arabidopsis thaliana (thale cress).  
 ACCESSION  
 VERSION  
 KEYWORDS  
 SOURCE  
 ORGANISM  
 Arabidopsis thaliana (thale cress)  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsids.  
 1 (bases 1 to 1143)  
 Castelli,V., Aury,J.M., Jaillon,O., Wincker,P., Clepet,C., Menard,M., Cruaud,C., Queller,F., Scarpelli,C., Schachter,V., Temple,G., Caboche,M., Weissenbach,J. and Salanoubat M.  
 Whole Genome Sequence Comparisons and 'Full-length' cDNA Sequences: A Combined Approach to Evaluate and Improve Arabidopsis Genome Annotation  
 Unpublished  
 2 (bases 1 to 1143)  
 Genoscope.  
 Direct Submission  
 Submitted (18-NOV-2003) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr)  
 - Web : www.genoscope.cns.fr)  
 The sequences are based on single pass reads.  
 Life Technologies (a division of Invitrogen) members carried out full-length libraries construction : Temple G.  
 Genoscope members carried out sequencing and annotation : Castelli V., Aury J.M., Jaillon O., Wincker P., Menard M., Cruaud C., Schachter V., Weissenbach J., Salanoubat M.  
 URGV INRA : Clepet C., Caboche M.  
 Annotation is based on the June 2003 version of the Arabidopsis genome released by MIPS (Munich Information center for Protein Sequences). 5 prime and 3 prime are assembled with Phrap.  
 http://www.genoscope.cns.fr/externe/sequences/Banque\_Projetc\_Fr/Full\_length  
 http://www.genoscope.cns.fr/cgi-bin/ggb/ggb?source=Arabidopsis.  
 FEATURES  
 source  
 1..1143  
 /organism="Arabidopsis thaliana"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:3702"  
 /clone="GSLTFB30ZG08"  
 /tissue\_type="flowers and buds"  
 /ecotype="Col-0"  
 /plasmid="pCMVSPORT 6"  
 complement(1..1143)  
 /gene="At5g5640"  
 ORIGIN  
 Alignment Scores: 0.335 Length: 1143  
 Pred. No.: 111.00 Matches: 58  
 Score: 37.3% Conservative: 23  
 Percent Similarity:





**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: January 20, 2006, 14:10:00 ; Search time 1091 Seconds  
(without alignments)  
4489.962 Million cell updates/sec

Title: US-10-783-417-2  
Perfect score: 3869  
Sequence: 1 MNQNNNDNNVEYIIIDSHSPY.....KLETIQTKNPFTNHTKTL 735

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 4996997 seqs, 3332346308 residues  
Total number of hits satisfying chosen parameters: 9993994

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:

-MODEL=frame+gdn.model -DEV=Xlh  
-O=/cgn2.1/USPTO\_epool/US10783417/runat\_20012006\_095009\_23859/app\_query.fasta\_1.903  
-DB=N Geneseg -OFMT=fastcap -SUFFIX=p2n.rng -MINMATCH=0.1 -LOOPEXT=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=blonum62 -TRANS=human40.cdi -LIST=45  
-DOCLIGN=200 -THR\_SCORE=pcr -THR\_MAX=100 -THR\_MIN=0 -MODE=LOCAL  
-OUTFMT=ptco -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US10783417@cgn\_1\_1\_752 @runat\_20012006\_095009\_23859 -NCPU=6 -ICPU=3  
-NO MMAP -NEG SCORES=0 -WAIT -DSPBLOCK=100 -DEV TIMEOUT=120 -WARN TIMEOUT=30  
-THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7 -YGAPOP=10  
-YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

N\_Geneseg\_21.\*  
1: genesegm19808.\*  
2: genesegm19908.\*  
3: genesegm20008.\*  
4: genesegm20018.\*  
5: genesegm20018.\*  
6: genesegm20028.\*  
7: genesegm20028.\*  
8: genesegm20038.\*  
9: genesegm20038.\*  
10: genesegm20038.\*  
11: genesegm20038.\*  
12: genesegm20048.\*  
13: genesegm20048.\*  
14: genesegm20058.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3869	100.0	2208	13	ADR89394 AXMI-006
2	3314	85.7	2235	13	ADR89396 AXMI-007
3	3104.5	80.2	2085	13	ADR89398
4	1061.5	27.4	4571	1	AAN93059 Delta-end

5	1060.5	27.4	3543	2	AAQ14669	AaQ14669 Dipteran
6	1060.5	27.4	3543	2	AAQ81178	AaQ81178 B.c. toxi
7	1059.5	27.4	3940	2	AAN93054	Aan93054 Delta-end
8	1059.5	27.4	4934	1	AAN81490	Aan81490 Insectici
9	1032.5	26.7	2130	12	ADP71294	AdP71294 Codon-mod
10	1032.5	26.7	2246	12	ADP71295	AdP71295 Codon-mod
11	899.5	23.2	2061	2	AAQ14670	AaQ14670 Dipteran
12	899.5	23.2	2061	2	AAQ81180	AaQ81180 B.c. toxi
13	881	22.8	3756	1	AAN50525	Aan50525 Bacillus
14	849	21.9	4896	6	AAQ43974	AaQ43974 Bacillus
15	849	21.9	4896	10	ADP31301	AdP31301 Bacillus
16	849	21.9	4896	10	ADP31306	AdP31306 Bacillus
17	775.5	20.0	3507	2	AAT43221	Aat43221 Antiscara
18	759	19.6	3684	6	ABK14949	AbK14949 Bacillus
19	759	19.6	3684	14	AEA61392	Aea61392 Bacillus
20	758	19.6	3684	1	AAN82106	Aan82106 Sequence
21	757.5	19.6	3535	1	AAN91003	Aan91003 Sequence
22	747.5	19.3	4184	1	AAN96139	Aan96139 Bacillus
23	711.5	18.4	4065	1	AAN92515	Aan92515 Sequence
24	705	18.2	4391	14	AEA61394	Aea61394 Bacillus
25	703	18.2	3633	6	ABK87235	AbK87235 Bacillus
26	703	18.2	3633	12	ADL15306	AdL15306 B. thuring
27	703	18.2	3633	14	ABE45608	AbE45608 B. thurin
28	703	18.2	6613	6	ABK87248	AbK87248 Bacillus
29	703	18.2	6613	12	ADL15321	AdL15321 B. thuring
30	703	18.2	6613	14	ABE45623	AbE45623 B. thurin
31	696.5	18.0	10974	1	AAN60055	Aan60055 VB131 clo
32	689.5	17.8	4188	8	ABK13487	AbK13487 B. popill
33	689.5	17.8	4188	14	ABE90816	AbE90816 Paenibaci
34	678.5	17.5	2003	6	ABK87237	AbK87237 Bacillus
35	678.5	17.5	2022	6	ABK87242	AbK87242 Bacillus
36	678.5	17.5	2022	12	ADL15316	AdL15316 B. thuring
37	678.5	17.5	2022	14	ABE45618	AbE45618 B. thurin
38	677	17.5	2478	12	ADP43079	AdP43079 Bacillus
39	676	17.5	1953	2	AAT43223	Aat43223 Antiscara
40	676	17.5	1953	3	AAQ09499	AaQ09499 Plant-cop
41	676	17.5	2425	2	AAQ05679	AaQ05679 bPGS1208
42	675	17.4	3621	6	ABK87234	AbK87234 Bacillus
43	675	17.4	3621	12	ADL15304	AdL15304 B. thuring
44	675	17.4	3621	14	ABE45606	AbE45606 B. thurin
45	675	17.4	3621	14	ABE34684	AbE34684 Bacillus

#### ALIGNMENTS

RESULT 1	
ADR89394	
ID	ADR89394 standard; cDNA; 2208 BP.
XX	
AC	ADR89394;
XX	
DT	18-NOV-2004 (first entry)
XX	
DE	AXMI-006 coding sequence.
XX	
KW	ss; gene; delta-endotoxin; delta-endotoxin associate polypeptide;
KW	expression cassette; transfection; transgenic plant; bacteria;
KW	lepidoptera; coleoptera; pest; pesticide; resistance;
KW	pesticidal activity.
XX	
OS	Bacillus thuringiensis.
XX	
FT	Key
FT	CDS
FT	FT
FT	FT
XX	1..2208
XX	/tag= a
XX	/product= "AXMI-006"
XX	
XX	WO2004074462-A2.
XX	
PD	02-SEP-2004.
XX	
XX	20-FEB-2004; 2004WO-US005829.
PF	
XX	



```

Db      1381 ATAAATTAATCAAAATGAACCTTATTAAATGCTCATCAACACACTCAAAATATCA
Qy      481 AAlaGlySerLeuSerAsnTyrGlnAsnThrPhePheGlnPheProArgLys
Db      1441 GCAGAGGGCTCTTAATCTAATATCAAAACAACTTTTTCATTTCTTGAAAANA
Qy      501 AspCysAsnLeuValIleAspProGlyCysSerProAsnAsnTyrSerHisIle
Db      1501 GACTGCATCTAATGATTAATGATCAGTGTTCACCAAACTTTAATATGATGATAT
Qy      521 LeuSerHisPheSerLeuPheThrTyrSerTyrValIleGlyLeuGlnIleLeu
Db      1561 TTATCCCATTTTTCATTTACTATTCTTATTCCTATGATGATGATCAGTCAAAATATTA
Qy      541 AspThrGlyValLeuGlyTyrThrHisSerValAspArgTyrAsnAlaIleSerAsp
Db      1621 GATACAGAGTGATTAAGATGAGACACAGTACTGTGATGATTAATGCAATATCAAT
Qy      561 LysIleIleThrMetIleProAlaIleLysGlyAsnAsnLeuAspThrAsnSerLysVal
Db      1681 AAAAATTAATCAATATCCACGATCAAAAGTAAATCTTGATACAAACTCTAAGTA
Qy      581 IleGlyGlyProGlyHisThrGlyValAsnLeuValTyrLeuGlnSerGlnGlyArgLeu
Db      1741 ATTGAAGAGCCTGTCATACAGAGAAACTGTTTATTATTAACAAGTCAAGGCGTTTA
Qy      601 GluIleThrCysGluThrProAsnSerThrGlnSerTyrPheIleArgLeuArgTyrAla
Db      1801 GAAATTAATCAATGTAAGTCTCTAATCTCAACATCTTATTTATGATGATGATGATGAT
Qy      621 ThrAsnGlyAlaGlyAsnThrLeuProAsnIleSerLeuThrIleProGlyValIleGly
Db      1861 ACAATGAGTCTGGAATATCTCTCTAATATATCTTACAAATACAGAGTAAATAGCA
Qy      641 IleProGlnArgLeuAsnAsnThrPheSerGlyThrAsnTyrAsnAsnLeuGlnTyr
Db      1921 ATACACCTCAACGACTCAACAACTTTTCTGATCAAAATTAATATTAATTAACAATAC
Qy      661 GlyAspPheGlyTyrPheGlnPheProSerThrValIleThrLeuProLeuAsnArgAsnIle
Db      1981 GGAGATTTTGGGTATTTCCAAATTTCCAAATGACAGTACATCTTCTTAAATGAAAACATA
Qy      681 ProPheIlePheAsnArgAlaAspValSerAsnSerIleLeuIleIleAspLysIleGlu
Db      2041 CCATTAAATTAATTAATCGACATGATATCAAAATTTAATATGATGATTAATAAATGAA
Qy      701 PheIleProIleThrSerSerMetHisGlnAsnArgGluLysGlnLysLeuGlnThrIle
Db      2101 TTTATACCAATTAATCTCTCTATGCACCAAATAGGAAAAACAAAAATTAAGAAACTATC
Qy      721 GlnThrLysIleAsnThrPhePheThrAsnHisThrLysThrLeu
Db      2161 CAAACAAAATAATATACATTTTTCACAAATCATACAAAACACTT 2205

```

```

RESULT 2
ADDR89396 ID ADDR89396 standard; cDNA; 2235 BP.
XX AC ADDR89396;
XX DT 18-NOV-2004 (first entry)
XX DE AXMI-007 coding sequence.
XX ss: gene; delta-endotoxin; delta-endotoxin associate polypeptide;
XX expression cassette; transformation; transgenic; plant; bacteria;
XX lepidoptera; coleoptera; pest; pesticide; resistance;
XX KM pesticidal activity.
XX OS Bacillus thuringiensis.
XX FH Key Location/Qualifiers

```

```

FT CDS 1..2235
FT /tag= a
FT /product= "AXMI-007"
FT /trans_except= pos:1..3, aa:Met
FN MO2004074462-A2.
PD 02-SEP-2004.
XX 20-FEB-2004; 2004MO-US005829.
XX
PR 20-FEB-2003; 2003US-0448632P.
PR 20-FEB-2003; 2003US-0448633P.
PR 20-FEB-2003; 2003US-0448797P.
PR 20-FEB-2003; 2003US-0448806P.
PR 20-FEB-2003; 2003US-0448810P.
PR 20-FEB-2003; 2003US-0448812P.
PR 19-FEB-2004; 2004US-00781979.
PR 19-FEB-2004; 2004US-00782020.
PR 19-FEB-2004; 2004US-00782096.
PR 19-FEB-2004; 2004US-00782141.
PR 19-FEB-2004; 2004US-00782570.
PR 19-FEB-2004; 2004US-00783417.
XX
PA (ATHE-) ATHENIX CORP.
PI Carozzi N, Hargis T, Koziel MG, Duck NB, Carr B;
XX WPI; 2004-635574/61.
DR P-PSDB; ADR89397.
XX
PT New isolated delta-endotoxin and delta-endotoxin-associated nucleic acids
PT and polypeptides; useful for killing lepidopteran or coleopteran pests or
PT for producing organisms with pesticide resistance.
XX
PS Claim 1; SEQ ID NO 8; 178bp; English.
XX
CC This sequence encodes an isolated delta-endotoxin. Some of the delta-
CC endotoxin coding sequences of the invention have alternative start
CC codons, producing more than one protein from a single open reading frame.
CC The nucleic acid sequences of the invention are useful in DNA constructs
CC or expression cassettes for transformation and expression in plants and
CC bacteria. The nucleic acids and corresponding polypeptides are useful for
CC killing lepidopteran or coleopteran pests. Compositions containing the
CC delta-endotoxins of the invention, and methods for their production, are
CC useful for the production of organisms with pesticide resistance,
CC specifically bacteria and plants. These organisms are useful for
CC generating altered or improved delta-endotoxin or delta-endotoxin-
CC associated proteins that have pesticidal activity, or for detecting the
CC presence of delta-endotoxin or delta-endotoxin-associated proteins or
CC nucleic acids in products or organisms.
XX
SQ Sequence 2235 BP; 861 A; 364 C; 316 G; 694 T; 0 U; 0 Other;
XX

```

```

Alignment Scores:
Pred. No.: 9, 256-290 Length: 2235
Score: 3314.00 Matches: 638
Percent Similarity: 91.3% Conservative: 32
Best Local Similarity: 86.9% Mismatches: 60
Query Match: 85.7% Indels: 4
DB: 13 Gaps: 3

```

```

US-10-783-417-2 (1-735) x ADR89396 (1-2235)
Qy 1 MetAENGInAsnAspAspAsnGlnGluIleIleAspSerHisThrSerProTyr 20
Db 1 GTGATCAAAATTAAT---AATATGATATATGATTAATCGATTCAAAGAAATTATCTAT 57
Qy 21 PheProAsnArgAsnSerAsnAspSerArgTyrProTyrThrAsnAsnProAsnGlnPro 40
Db 58 CTTTCTAACAGAAATTAATGATCATCTTATGATACCCCTTACACAAATATTCGAATCAACCA 117
Qy 41 LeuGlnAsnThrAsnTyrLysGluThrPheLeuAsnMetCysGlnGlyAsnThrGlnTyrGly 60

```

```
Db      118 TTACAAAACAAAATTAACAAAGAGTGCTCAATATGTGCAGGGAATACAAATATGCT 177
Qy      61 AspaenPhegluThrPheAlaSerAlaAspThrIleAlaValSerAlaGlyThrIle 80
Db      178 GATATTTTCGACATCTTGCTAGTGTGATACAAATTCGTCGAGTTAGTGCAGTACTATT 237
Qy      81 ValSerGlyThrLeuLeuAlaGlyIleGlyGlyLeuThrSerIleSerGlyProIleGly 100
Db      238 GTATCCGGTACTCTGTAGCCGGTAAAGTGGGCTCACTTCTATATCCGGAACCATAGGA 297
Qy      101 IleIleGlyAlaIleIleIleSerPheGlyThrIleIleThrValPheThrProAlaGly 120
Db      298 ATAAATAGGTGCTATATATATATCTTTGTGATCCCAATACATGCTTTGGCCCGGGA 357
Qy      121 GluGlnAspLysThrValTyrThrGlnPheIleLysMetGlyGluIlePheValAspThr 140
Db      358 GAAACAAGACAAACAGTATGACACAAATTTATTTAAATGGAGAAATTTTGTGATACA 417
Qy      141 ProLeuThrGluSerIleLysGlnLeuLysLeuGlnThrLeuGluGlyPheArgGlnIle 160
Db      418 CCGTTAAACGAAAGCATAAACAGCTAAAGTTACAAACTTTAGAAAGATTTAGACAAATA 477
Qy      161 LeuGlnSerTyrAsnThrAlaLeuAspAspTyrParglyLysLeuLysArgLeuGlnAlaPro 180
Db      478 TTACAAAGCTATAATACAGCATTAATGATGATGAGAAATTTAAAGACTACAAAGCTCCT 537
Qy      181 GlyLeuProPheSerSerAlaLeuGlnGlnAlaAlaLeuThrLeuLysIleArgPheGly 200
Db      538 GATTAACCAACATTCAGCATTTACACAAAGCTGCTTGACTTTAAATACGATTTGAG 597
Qy      201 AsnValHisAsnAspPheIleArgGlnIleProGlyPheGlnLeuGlnThrTyrLysThr 220
Db      598 AATGTTCAACAATGATTTTATTCAGAAATATCCTGGTTTCAACTTGAAACTTATAAACG 657
Qy      221 LeuLeuLeuProIleTyrAlaGlnAlaAlaAsnPheHisLeuAsnLeuLeuGlnGlnGly 240
Db      658 CTATTACTACTAATTTATATGCGCAAGCTGCTAATTTTCATTTAAATTTATCAACAAGCT 717
Qy      241 AlaGluLeuAlaAspGluTyrAsnAlaAspIleHisProSerGlnIleGluProAsnAla 260
Db      718 GCTGAATGGCTGATGAATGGAATGCAATATACATCTTCAACAATGAACTTAATGCT 777
Qy      261 GlyThrSerAspAspTyrTyrLysLeuLeuLysGlnAlaIleProLysTyrSerAsnTyr 280
Db      778 GGAACATCAGATGACATATTAATACTTTAAAGAAAATATACCTTAATATGTAACAT 837
Qy      281 CysAlaAsnThrTyrArgThrGlyLeuLysAsnLeuArgAspGluProAsnMetLysTyr 300
Db      838 TGTGCAAAATACCTATAGAGAAAGACCTAAATAAATTCGAAACGAACTAATATGAGATGG 897
Qy      301 SerIlePheAsnAspTyrArgArgTyrMetThrIleThrValLeuAspThrIleSerGln 320
Db      898 AGTAAATATTAATGATTAATGAGATATATGACTATTAATGATTAATGATTAATGATTAAT 957
Qy      321 PheSerLeuTyrAspIleLysArgTyrArgAspSerIleGlyGlyIleGluValLysGly 340
Db      958 TTTTCTTTTATGATTAATAGATACAAAGATTCATAGGA-----AGAAATAGTGGC 1011
Qy      341 IleLysAsnGluLeuThrArgGluIleTyrThrThrGluIleAsnPheAspArgLeuPro 360
Db      1012 ATTTAAACGAACTTACAAAGAAATTTATACACAGAAATTAATTTTACACGCTCTTACT 1071
Qy      361 GlnLeuArgValGlnProAsnLeuAlaThrMetGluTyrAsnLeuThrArgAlaSerPhe 380
Db      1072 TACCTTGAATTCACCCCAATCTCGCTAATATGAGATATTAATTAACAGTTCAGAGGCTT 1131
Qy      381 LysLeuPheSerPheLeuGlnGlnPheIlePheTyrThrGluAsnThrAsnPheGlnVal 400
Db      1132 ACATTAATTTTCAATTTTATGATGACATTAATTTTATACAAATAAGAAACGTACGGGAAT 1191
Qy      401 ArgLeuValGlyIleSerAsnArgAspAlaProThrTyrSerAsnThrIleThrGluThr 420
Db      1192 CGTTTAGTGGTATTTGCGAATCGTAATGATCTACTTATGCTAGCAGCAGACGTAATTT 1251
Qy      421 LeuTyrGlyValArgThrGlySerProThrThrLysThrIleArgProPheGluSerTyr 440
Db      1252 ATATATGAGAAAGAAACAGGTCCACCCACAAACAAATTTAATATACATTTTGAATCTCAT 1311
Qy      441 LysValSerIleValThrAspArgGlnSerProProValSerProIle--GlnProHis 459
Db      1312 AAAGTTTCAATTTGTAACGATAGACAAAGTAACCTCTTCCCTTTCCTTAACATATAC 1371
Qy      460 PheIleIleAsnGlnIleGluLeuTyrLeuAsnGlySerSerAsnAsnThrLeuLysTyr 479
Db      1372 TTACAAATTAATCAAAATTTAACTTTATTTAATTAATTCACCTAGTAATTAATTAACATAT 1431
Qy      480 SerAlaGlySerLeuSerSerAsnTyrGlnAsnThrThrPhePheGlnPheProArgLys 499
Db      1432 TCAGCTGGGGGGAATTTATCTAAATGATTAATAAAACAACTGATTTTCAATTTCTGTAAA 1491
Qy      500 LysAspCysAsnLeuValIleAspProGlyCysSerProAsnPheAsnAsnTyrSerHis 519
Db      1492 AAAGACTGAACCAATTAATTAATCCAAATTTGTTACCAAGCTAATTAATGTTATATGCTAT 1551
Qy      520 IleLeuSerHisPheSerLeuPheThrTyrSerTyrValIleGlyLeuGlnLeuGlnIle 539
Db      1552 ATTTATCCAGTTTCTTATTTAATTAATTAATTTCTTAATAAATTTGATTAATGCGCTAATATA 1611
Qy      540 LeuAspThrGlyValIleGlyTyrThrHisSerSerValAspArgTyrAsnAlaIleSer 559
Db      1612 TTATATACAGGTGCATTAATGATGACACACAGTATGTTAATTAAGAAATTAATGCAATATCA 1671
Qy      560 AspLysIleIleThrMetIleProAlaIleLysGlyAsnAsnLeuAspThrAsnSerLys 579
Db      1672 GATTAATAATTAATTAATCAATATCCAGCATCAAGGTAAACAGTCTTGAATCAAACTCTAAG 1731
Qy      580 ValIleGluGlyProGlyHisThrGlyLysAsnLeuValTyrLeuGlnSerGlnGlyArg 599
Db      1732 GTAATTAAGAACCTGCTGCATACAGAGAGAAACTTGTTTATTTTCAAAAGTCAAAGGCGCT 1791
Qy      600 LeuGluIleThrCysGluThrProAsnSerThrGlnSerTyrPheIleArgLeuArgTyr 619
Db      1792 TTACAGATTAATGATGAGAACTCCATATTTCTACACATTTTATTAATTAATTAATTAATTA 1851
Qy      620 AlaThrAsnGlyAlaGlyAsnThrLeuProAsnIleSerLeuThrIleProGlyValIle 639
Db      1852 GCTCAAAATGGGTGCGAAATATCTCTCTATATATCTCTTACAAATACAGAGATATA 1911
Qy      640 GlyIleProGlnArgLeuAsnAsnThrPheSerGlyThrAsnTyrAsnAsnLeuGln 659
Db      1912 GGAATACCACTCAACAGCTCAACAAACCTTTTCTGTACCAAAATTAATTAATTAATTAATTA 1971
Qy      660 TyrGlyAspPheGlyTyrPheGlnPheProSerThrValThrLeuProLeuAsnArgAsn 679
Db      1972 TACGGAATTTTGGTATTTTCAATTTTCAAGTACAGTAACATTTTCAATTAATTAATTAATTA 2031
Qy      680 IleProPheIlePheAsnArgAlaAspValSerAsnSerIleLeuIleIleAspLysIle 699
Db      2032 ATACATTAATTAATTAATTCGAGAGATGATACAAATTCATTAATTAATTAATTAATTAATTA 2091
Qy      700 GluPheIleProIleThrSerSerMetHisGlnAsnArgGluLysGlnLysLeuGluThr 719
Db      2092 GAATTTAATCAAAATTAATCTCTCTCTGTCGCAAAATTAAGAAACAAATAATTAAGAAACT 2151
Qy      720 IleGlnThrLysIleAsnThrPhePheThrAsnHisThrLys 733
Db      2152 ATCCAAACAAATAATTAATTAATTTTCAAAATCATACAAATA 2193
Db      18-NOV-2004 (first entry)
```

XX AXMI-007 alternative start site coding sequence.  
DE ss; gene; delta-endotoxin; delta-endotoxin associate polypeptide;  
XX expression cassette; transformation; transgenic; plant; bacteria;  
KW lepidoptera; coleoptera; pest; pesticide; resistance;  
XX pesticidal activity.  
OS *Bacillus thuringiensis*.  
XX  
XX Key Location/Qualifiers  
FT CDS 1..2085  
FT /tag= a  
FT /product= "Alternative AXMI-007"  
PN WO2004074462-A2.  
XX  
XX 02-SEP-2004.  
PD 20-FEB-2004; 2004WO-US005829.  
XX  
XX 20-FEB-2003; 2003US-0448632P.  
PR 20-FEB-2003; 2003US-0448632P.  
PR 20-FEB-2003; 2003US-0448633P.  
PR 20-FEB-2003; 2003US-0448797P.  
PR 20-FEB-2003; 2003US-0448806P.  
PR 20-FEB-2003; 2003US-0448810P.  
PR 20-FEB-2003; 2003US-0448812P.  
PR 19-FEB-2004; 2004US-00781979.  
PR 19-FEB-2004; 2004US-00782020.  
PR 19-FEB-2004; 2004US-00782096.  
PR 19-FEB-2004; 2004US-00782141.  
PR 19-FEB-2004; 2004US-00782570.  
PR 19-FEB-2004; 2004US-00783417.  
XX  
XX (ATHE-) ATHENIX CORP.  
PA  
PI Carozzi N, Hargies T, Koziel MG, Duck NB, Carr B;  
XX WPI: 2004-635574/61.  
DR P-PSDB; ADR89399.  
XX  
XX New isolated delta-endotoxin and delta-endotoxin-associated nucleic acids  
PT and polypeptides, useful for killing lepidopteran or coleopteran pests or  
PT for producing organisms with pesticide resistance.  
XX  
XX Claim 1; SEQ ID NO 10; 178bp; English.  
XX  
XX This sequence encodes an isolated delta-endotoxin. Some of the delta-  
CC endotoxin coding sequences of the invention have alternative start  
CC codons, producing more than one protein from a single open reading frame.  
CC The nucleic acid sequences of the invention are useful in DNA constructs  
CC or expression cassettes for transformation and expression in plants and  
CC bacteria. The nucleic acids and corresponding polypeptides are useful for  
CC killing lepidopteran or coleopteran pests. Compositions containing the  
CC delta-endotoxins of the invention, and methods for their production, are  
CC useful for the production of organisms with pesticide resistance,  
CC specifically bacteria and plants. These organisms are useful for  
CC generating altered or improved delta-endotoxin or delta-endotoxin-  
CC associated proteins that have pesticidal activity, or for detecting the  
CC presence of delta-endotoxin or delta-endotoxin-associated proteins or  
CC nucleic acids in products or organisms.  
XX  
XX Sequence 2085 BP; 794 A; 338 C; 302 G; 651 T; 0 U; 0 Other;  
SQ

Alignment Scores:  
Pred. No.: 7.9e-271 Length: 2085  
Score: 3104.50 Matches: 598  
Percent Similarity: 91.9% Conservative: 30  
Best Local Similarity: 87.6% Mismatches: 52  
Query Match: 80.2% Indels: 3  
DB: 13 Gaps: 2

US-10-783-417-2 (1-735) x ADR89398 (1-2085)

QY 52 MetCysGInGlyAenThrGlnTyrGlyAspAsnPheGluThrPheAlaSerAlaAspThr 71  
DB 1 ATGTCTCAAGGAAATACACAAATATGATGATATTTCCAGACATTTGCTAGTCTATACA 60  
QY 72 ILeAlaAlaValSerAlaGlyThrIleValSerGlyThrLeuLeuAlaGlyIleGlyGly 91  
DB 61 ATTGCTGAGTTAGTACAGGTATCTATGATCCGGTACTCTGTATGCCGTATAGTGGG 120  
QY 92 LeuThrSerIleSerGlyProIleGlyIleIleGlyAlaIleIleIleSerPheGlyThr 111  
DB 121 CTCACCTCTAATATCCGAGCCGATAGAAATATAGGTGCTAATATATATCTTTTGCTACC 180  
QY 112 LeuIleThrValPheThrProAlaGlyGluGluAspLysThrValITrThrGlnPheIle 131  
DB 181 CTAATCAGCTGCTTTTGGCCGGGAGAAACAGACAAACAGTATGACACAAATTTATT 240  
QY 132 LysMetGlyGluIlePheValAspThrProLeuThrGlnSerIleIleGlnLeuLysLeu 151  
DB 241 AAAATGGAGAAATTTTGTGATATACCGTTAACAGAAAGCATTAAGCTTAAGTTA 300  
QY 152 GlnThrLeuGluGlyPheArgGlnIleLeuGlnSerTyrAsnThrAlaLeuAspAspTrp 171  
DB 301 CAACCTTATGAAAGATTTTATGACAAATATATACAACTATATACAGCATTTAGATGATGG 360  
QY 172 ArgLysLeuLysArgLeuGlnAlaProGlyLeuProProSerSerAlaLeuGlnGlnAla 191  
DB 361 AGAAATATTAAGATCAGCTCAGATTTACACATCTATCAGCATTTACACAGCT 420  
QY 192 AlaLeuThrLeuLysIleArgPheGluAsnValHisAsnAspPheIleArgGluIlePro 211  
DB 421 GCCTTGACTCTTAAATATGAGATTTGAGATGTTCAATGATTTTATTTCGAAATPACT 480  
QY 212 GlyPheGlnLeuGluThrTyrLysThrLeuLeuLeuProIleTyrAlaGlnAlaAsn 231  
DB 481 GGTTCCAACTGAACTTAATTAACGCTATTAATTAATTAATTAATTAATTAATTAAT 540  
QY 232 PheHisLeuAsnLeuLeuGlnGlnGlyAlaGluLeuAlaAspGluITrPheAsnIleAspIle 251  
DB 541 TTTCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 600  
QY 252 HisProSerGlnIleGlnProAsnAlaGlyThrSerAspAspTyrTyrLysLeuLeuLys 271  
DB 601 CATCTTCACAAATTTGAACCTAATCTGGAACATCAGATGACTATTATTAATCTTTTAAA 660  
QY 272 GluAsnIleProLysTyrSerAsnTyrCysAlaAsnThrTyrArgThrGlyLeuLysAsn 291  
DB 661 GAAATATATACCTAAT 720  
QY 292 LeuArgAspGluProAsnMetLysITrSerIlePheAsnAspTyrArgArgTyrMetThr 311  
DB 721 CTTCCAAACGAACTTAATATATGAGATATATTAATTAATTAATTAATTAATTAATTA 780  
QY 312 IleThrValLeuAspThrIleSerGlnPheSerLeuTyrAspIleLysArgTyrArgAsp 331  
DB 781 ATTACTGATTTATGATCATCTCATATTTCTTTTATGATATATTAAGATATCAAAAGT 840  
QY 332 SerIleGlyGlyIleGluValIleLysValIleLysAsnGluLeuThrArgGluIleTyrThr 351  
DB 841 TCAATATAGG-----AGAAATAGTGCATTTAAACGAACTTAACAGAAATTTATACA 894  
QY 352 ThrGluIleAsnPheAspArgLeuProGlnLeuAlaGlnPheAsnLeuAlaThrMet 371  
DB 895 ACTGAAATTAATTTTGGACCGTCTTACTTCACTGAATTAACCAACCAATCTCCCTATTAAG 954  
QY 372 GluTyrAsnLeuThrArgAlaSerPheLysLeuPheSerPheLeuGlnGlnPheIlePhe 391  
DB 955 GAAATATATTTATACAGTTCAGGCGCTTATTTTATTTTATTTTATTTTATTTTATTT 1014  
QY 392 TyrThrGluAsnThrAsnPheGlyAsnArgLeuValIleGlyIleSerAsnArgAspAlaPro 411  
DB 1015 TATACAAAAAATGAACGATACGGAATGCTTATGATTCGAAATCTGATTAATGATCT 1074

```

QY 412 ThrTyrSerAsnThrIleTherGluThrLeuTyrGlyGluArgThrGlySerProThrThr 431
Db 1075 ACTTATGCTACGACGAGAACTTAAATATATATGAGAAAGAGGCTCCACCAACA 1134
QY 432 LysThrIleArgProPheGluSerTyrIleValSerIleValThrAspArgGlnSerPro 451
Db 1135 AAAACCTTAAATCCATTGAACTCTTAATAGTTTCAATTGACTGATAGACAGTAAGT 1194
QY 452 ProValSerProIle--GlnProHisPheIleIleAsnGlnIleGluLeuTyrLeuAsn 470
Db 1195 CCTACTCCCTCTTCTCTAAACATATACCTTACAAATTAATCAATTAATGAATTTAATTAAT 1254
QY 471 GlySerSerAsnAsnThrLeuLysTyrSerIleArgIleGlySerLeuSerAsnTyrGlnAsn 490
Db 1255 AATTACCTAGTAATTAATTAATCAATTAATCAATGCGGGGGAATTTATCTAATGATGATAA 1314
QY 491 ThrThrPhePheGlnPheProArgIleLysAspCysAsnLeuValIleAspProGlyCys 510
Db 1315 ACNACTGATTTTCAATTTCTCTTAAAAAAGACGTGTAACCAATTAATTAATCAATTTGT 1374
QY 511 SerProAsnPheAsnAsnTyrSerHisIleLeuSerHisPheSerLeuPheThrTyrSer 530
Db 1375 TTACCAAGCTAATATAGTATATAGTATATTTATCCAGTTTCTTTATTTAATTAATTC 1434
QY 531 TyrValIleGlyLeuGlnLeuGlnIleLeuAspThrGlyValLeuGlyTyrThrHisSer 550
Db 1435 TATATAATTTGATGTCGCTAATATATATATATATATATATATATATATATATATATATAT 1494
QY 551 SerValAspArgTyrAsnAlaIleSerAspLysIleIleTherMetIleProAlaIleLys 570
Db 1495 AGTGTATATAGAAATATATGCAATATACCAATTAATTAATTAATTAATTAATTAATTAAT 1554
QY 571 GlyAsnAsnLeuAspThrAsnSerLysValIleGluGlyProGlyHisThrGlyGlyAsn 590
Db 1555 GGTAACAGCTCTGATACAACTCTAAGGTAATGAGAAGCCTGCTCATACAGAGAAAC 1614
QY 591 LeuValIleGlyLeuGlnSerGlnIleGlyLeuGlnIleThrCysGluThrProAsnSerThr 610
Db 1615 TTGGTTTATTTACAAAGTCAAAGGCGGTTAGATTAATGATGAACTCTTAATTTCTACA 1674
QY 611 GlnSerTyrPheIleArgLeuArgTyrAlaThrAsnGlyAlaGlyAsnThrLeuProAsn 630
Db 1675 CAATCTTATACATTAGACTTGCATACGCTACAAATGGTCTGGAATATCTCTTCTAAT 1734
QY 631 IleSerLeuThrIleProGlyValIleGlyIleProProGlnArgLeuAsnAsnThrPhe 650
Db 1735 ATATCTCTTACAAATCCAGAGTATAGGAATACCACTCAACAGCTCAACAGACTTTT 1794
QY 651 SerGlyThrAsnTyrAsnAsnLeuGlnTyrGlyAspPheGlyTyrPheGlnPheProSer 670
Db 1795 TCTGGTACAAATTAATTAATTAATTTACAAATACGAGATTTTGGGTATTTCCAAATTTCCAAGT 1854
QY 671 ThrValThrLeuProLeuAsnArgAsnIleProPheIlePheAsnArgAlaAspValSer 690
Db 1855 ACAGTAACATTAACCTTAAATGAAACATACCATTTATATTTAATCGAGATGATATCA 1914
QY 691 AsnSerIleLeuIleIleAspLysIleGluPheIleProIleThrSerSerMetHisGln 710
Db 1915 AATTCATTTTATATCATTTGATTAATTTGATTAATTTAATTAATTAATTAATTAATTAAT 1974
QY 711 AsnArgGlyLysGlnLysLeuGlnThrIleGlnThrLysIleAsnThrPhePheThrAsn 730
Db 1975 AATTAAGAAAAACAAAAATTAGAACTATCCAAACAAAAATTAATATATATTTTTCACAAAT 2034
QY 731 HisThrLys 733
Db 2035 CATACAAA 2043

```

RESULT 4  
ID AAN93059 standard; DNA; 4571 BP.  
XX  
AC AAN93059;

```

XX 27-AUG-2003 (revised)
DT 25-MAR-2003 (revised)
DT 30-MAY-1990 (first entry)
XX
DE Delta-endotoxin crystal protein gene.
XX
KM Delta-endotoxin; crystal protein; insecticide; pCH130; ss;
KM biological control agent.
OS
XX
XX Bacillus thuringiensis; israelensis.
XX
FH Key Location/Qualifiers
FT RBS /tag= b
FT /label= Shine-Dalgarno sequence.
FT misc_feature 891..4430
FT /tag= a
FT /product= "delta-endotoxin crystal protein"
XX
XX EP296870-A.
XX
XX
XX 28-DEC-1988.
XX
XX 24-JUN-1988; 88EP-00305772.
XX
XX 26-JUN-1987; 87US-00067653.
XX
XX (DUPLO ) DU PONT DE MEMOURS & CO E. I.
XX
XX Ellar DJ, Ward ES;
XX
XX WPI; 1989-001322/01.
XX
XX P-PSDB; AAP93715.
XX
XX
XX DNA fragment encoding insecticidal protein - obtd. from Bacillus
XX thuringiensis sub species israelensis, and used in microorganisms and
XX plant cells.
XX
XX Disclousure; Fig 7; 26pp; English.
XX
XX
XX The sequence encodes the 130 kDa delta-endotoxin gene from B.
XX thuringiensis subsp. israelensis and is a 4.46 kb insert of pCH130. The
XX delta endotoxin protein is insecticidal and can be used to control insect
XX pests esp. mosquitos. See also AAN93054. (Updated on 25-MAR-2003 to
XX correct PA field.) (Updated on 27-AUG-2003 to correct OS field.)
XX
SQ Sequence 4571 BP; 1678 A; 685 C; 817 G; 1391 T; 0 U; 0 Other;
Alignment Scores:
Pred. No.: 2.48e-85 Length: 4571
Score: 1061.50 Matches: 271
Percent Similarity: 50.4% Conservative: 118
Best Local Similarity: 35.1% Mismatches: 278
Query Match: 27.4% Indels: 105
Gaps: 24
US-10-783-417-2 (1-735) x AAN93059 (1-4571)
QY 1 MetAsnGlnAsnAsnAspAsnAsnGluTyrGluIleIleAspSerHisThrSerProTyr 20
Db 891 ATGATCTTATCAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 950
QY 21 PheProAsnArgAsnSerAsnAspSerArgTyrProTyrThrAsnAsnProAsnGlnPro 40
Db 951 -----AATATATCTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1004
QY 41 LeuGlnAsnThrAsnTyrLysGluTyrPheAsnMetCysGlnGlyAsnThrGlnTyrGly 60
Db 1005 TTACAAGTACAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1064
QY 61 AspAsnPheGluThrPheAlaSerAlaAspThrIleAlaIleValSerIleGlyThrIle 80

```



Db 1065 GGAGATTTTGAACCTTTATGTAAGT-----GGTGAACCTACGCTTAATCTATT 1115  
 QY ValSerGlyThrLeuLeuAlaGlyIleGlyGlyLeuThrSerIleSerGlyProIleGly 100  
 Db 1116 GTAGTTGGGACCGTACTGACTGCTTTGGGTTCAACA-----CCCTTAGGA 1163  
 QY 101 IleIleGlyAlaIleIleIleSerPheGlyThrLeuIleThrValPheProAlaGly 120  
 Db 1164 CTT-----GCTTAATAGGTTTGGTACATTAAACAGTCTTTTCCAGCCCA 1214  
 QY 121 GluGlnAspLysThrValIleThrGlnPheIleLysMetGlyGluIlePheValAspThr 140  
 Db 1215 GACCAATCTAACCA-----TGAGTGAAGCTTAATTAACAACAATAATTAATTAATAA 1271  
 QY 141 ProLeuThrGlnSerIleLysGlnLeuLysGlnThrLeuGlnIlePheArgGlnIle 160  
 Db 1272 GAAATGCAATCAACATTAATTAAGTAATGCTAATTAATTAATTAACAGTCTTAATGTT 1331  
 QY 161 LeuGlnSerThrAsnThrAlaLeuAspAspTrpArgLysLeuLysArgLeuGlnAlaPro 180  
 Db 1332 ATCAGACATTAATCATATACCTTAATAACATGGAG-----AATATCCA 1376  
 QY 181 GlyLeuProSerSerSerAlaLeuGlnAlaAlaLeuThrLeuLysIleArgPheGln 200  
 Db 1377 AACCCCAAAATATCAGATGTAAGACAAATCCAGCTAGTTCATTACCATTTTCA 1436  
 QY 201 AsnValHisAsnAspPheIleArgGluIlePro-----GlyPheGlnLeuGlnThr 218  
 Db 1437 AATGTCATTCAGAGCTGTAACTCTGTCTCTTAATCTAGTATTCAGATTAATCTAT 1496  
 QY 219 LysThrLeuLeuLeuProIleThrAlaGlnAlaAlaAsnPheHisLeuAsnLeuGln 238  
 Db 1497 AACATATCTGATTAATCTAGTATGTCACAGACGCAACTTACATCTGCTGATTAAT 1556  
 QY 239 GlnGlyAlaGluLeuAlaAspGluTrpAsnAlaAspIleHisProSerGlnIleGluPro 258  
 Db 1557 CAAGCGCTCAAAATTTGAGCGATTAATAAACAATCGAACATTCATTAATTAAGAGCT 1616  
 QY 259 AsnAlaGlyThrSerAspAspTrpTrpLysLeuLeuLysGluAsnIleProLysTrpSer 278  
 Db 1617 ---TTGCCAACAGCAATTAATTAATTCAGATTAATCACTAAGCTTAAGATTAACACT 1673  
 QY 279 AsnTrpCysAlaAsnThrTrpArgThrGlyLeuLysAsnLeuArgAspGluPro----- 296  
 Db 1674 AATTATTTGCTAACATTAATAAAGATTAATTAATTAATAAAGACGCTGATAGT 1733  
 QY 297 -----AsnMetLysTrpSerIlePheAsnAspTrpArgTrpMetThrIle 312  
 Db 1734 AATCTTGATGAAATTAATAACTGGAACACATCAATACGATCGAACAAATATGACTACT 1793  
 QY 313 ThrValLeuAspThrIleSerGlnPheSerLeuTrpAspIleLysArgTrpArgAspSer 332  
 Db 1794 GCTGATTTGATGTTGTCACCTCTTCTTAATTAATGATGATGATGATTAAT----- 1844  
 QY 333 IleGlyGlyIleGluValLysGlyIleLysAsnGlnLeuThrArgGluIleTrpThr 352  
 Db 1845 -----CCAAATAGGTGCAATCTGAATCTGACAAATTAATTAATCAAGTA 1889  
 QY 353 GluIleAsnPheAspArgLeuProGlnLeuArgValGlnProAsnLeuAlaThrMetGln 372  
 Db 1890 ---CTTAATCTGGAAGAAAGCCCTTAATAATTAAT-----GACTTCAATATCAAGAG 1940  
 QY 373 TyrAsnLeuThrArgAlaSerPheLysLeuPheSerPheLeuGlnGlnPheIlePheTrp 392  
 Db 1941 GATTCACCTTAACGTAAGCCG---CATTTATTTACTGCTGATTCCTTGAATTTTAT 1997  
 QY 393 -----ThrGluAsnThrAsnPheGlyAsnArg----- 401  
 Db 1998 GAAAAAGCGCAACCTACTTAATTAATTTTTCACCAAGCAATTAATAATGTTCACTATC 2057  
 QY 402 ---LeuValGlyIleSerAsnArgAspAlaProThrTrpSerAsnThrIleThrGluThr 420  
 Db 2058 AACCTTGATTAATATATCCAAAAAATCTAGTGTTTTGGAAATCACAATTAATGATTA 2117

QY 421 LeuTrpGlyGluArgThrGlySerProThrThrLysThrIleArgProPheGlnSerTrp 440  
 Db 2118 TTA-----AATCTCTGTTGGCAACAAATATTATTAAT----- 2153  
 QY 441 LysValSerIleValThrAspArgIleSerProValSerProIleGlnProHisPhe 460  
 Db 2154 -----TTT 2156  
 QY 461 IleIleAsnGlnIleGluLeu-----TyrLeuAsnGlySerSerAsnAsnThr--- 476  
 Db 2157 TTATTAATATGTCATAGCTAGATTAATAATCTTAATGATTAATAATTAATTTGATAA 2216  
 QY 477 -----LeuLysTrpSerAlaGly 483  
 Db 2217 ATGCAATTTTATTAATGTAATGTAATGTAATGTAATGTAATGTAATGTAATGTAAT 2276  
 QY 484 SerLeuSerAsnTrpGln---AsnThrThrPhePheGlnPheProArgLysAlaAspCys 502  
 Db 2277 GCGCAATTAATTAATGATGTAATTAATAATTAATTTCCGGGTACCAATCTTAACGAGA 2336  
 QY 503 AsnLeuValIleAspProGlyCysSerProAsnPheAsnAsnTrpSerHisIleLeuSer 522  
 Db 2337 GAGAACTCAAGAAACCTAACCTTTTCCACATGATGATGATGATGATGATGATGAT 2396  
 QY 523 HisPheSerLeuPheThrSerTrpValIleGlyLeuGlnLeuGlnIleLeuAspThr 542  
 Db 2397 TTTATTAATAAGCTTAATGTAATGTAATGTAATGTAATGTAATGTAATGTAATGTAAT 2444  
 QY 543 GlyValLeuGlyTrpThrHisSerSerValAspArgTrpAsnAlaIleSerAspLysIle 562  
 Db 2445 ---AGTTTGCCTGGACACACTGATGTTGATGTTGATGTTGATGTTGATGTTGATGTT 2501  
 QY 563 IleThrMetIleProAlaIleLysGlyAsnAsnLeuAspThrAsnSerLysValIleGlu 582  
 Db 2502 ACTACCAAAATTCAGCTGTAAAGCAATTCATGAGGAGCTGCTTAAGGTGTTCA 2561  
 QY 583 GlyProGlyHisThrGlyGlyLysLeuValTyrLeuGlnSerGlyArgLeuGluIle 602  
 Db 2562 GGAAGCTGGTCATACAGAGGGGATTAAT-----GATTCAAGATCATTTCAAAAT 2615  
 QY 603 ThrCysGluThrProAsnSerThrGlnSerTrpPheIleArgLeuArgTrpAlaThrAsn 622  
 Db 2616 ACATGTCACACACTCAAAATTTTCAACATGCTATTTTAAGAATTCGTTATCTTCAAT 2675  
 QY 623 GlyAlaGlyAsnThrLeuProAsnIleSerLeuThrIleProGlyValIleGlyIlePro 642  
 Db 2676 GGAACGCGAAATATCGAGCTGTAAATCTTAATATCCAGGGGTACAGAACTG--- 2732  
 QY 643 ProGlnArgLeuAsnAsnThrPheSerGlyThrAsnTrpAsnAsnLeuGlnTrpGlyAsp 662  
 Db 2733 GGTATGCACTAACCCCACTTTTCTGCTAGATTAATCAAGAAATTAATAATTAAGAT 2792  
 QY 663 PheGlyTrpPheGlnPheProSerThrValThrLeuProLeuAsnArgAsnIleProPhe 682  
 Db 2793 TTTCACTACTTAATTAATTTTCAACAGAGTAATTTGCTCCAAATCAAAATATCTCTT 2852  
 QY 683 IlePheAsnArgAlaAspVal---SerAsnSerIleLeuIleIleAspLysIleGluPhe 701  
 Db 2853 GTGTTTAATCGTTCCGATGATATATCAAAACAACAAGTACTTAATGATTAATTAATG 2912  
 QY 702 IleProIleThrSerSerMetHisGlnAsnArgGlyLysGlnLysLeuGlnIleGln 721  
 Db 2913 CTGCCAATTAATCTGTTCTATTAAGAAGATTAAGAGAAACAATAATTAAGAACAGTACA 2972  
 QY 722 ThrLysIleAsnThrPhePheThrAsnHisThrLys 733  
 Db 2973 CAAATTAATTAATCAATTTTATGCAATCTTAATAA 3008

RESULT 5  
 AAQ14669  
 ID AAQ14669 standard; DNA; 3543 BP.  
 XX



QY 421 LeuTYRGLYGLUARGThrGlySerProThrThyIleArgProPheGluSerTYR 440  
 DB 1228 TTA-----AAATCTCTGGTTGGCAACAATAATTATTAATT----- 1263  
 QY 441 LysValSerIleValThrAspArgInSerProValSerProIleGlnProHisPhe 460  
 DB 1264 -----TTT 1266  
 QY 461 IleIleGlnIleGluLeu-----TYRLeuGlnGlySerSerAsnThr--- 476  
 DB 1267 TTAATAAATGTCATAGCTTAGATATAATAATATCTAATAAGATTAATATATTGTGATAA 1326  
 QY 477 -----LeuTYRSerIleGlyGly 483  
 DB 1327 ATGATTTTATATACTAATAGTACTAGACTTTGGAGAAAGAACTTACAGAGGATCT 1386  
 QY 484 SerLeuSerAsnTYRGLN---AsnThrThrPhePheGlnPheProArgLysLysAspCys 502  
 DB 1387 GGGCAATAAATCTTAGATGATTAATAATAATATTTTCGGTTCCCAATTCTTAACCAACA 1446  
 QY 503 AsnLeuValIleAspProGlyCysSerProAsnPheAsnAsnTYRSerHisIleLeuSer 522  
 DB 1447 GAGAAATCAAGCAATCCCTACCTTTTCCACATATGATTAATAGTATATTTATCA 1506  
 QY 523 HisPheSerLeuPheThrTYRserTYRValIleGlyLeuGlnLeuGlnIleLeuAspThr 542  
 DB 1507 TTTATTTAAAGTCTTAGTATCCCTGCACATATAAATCAAGTAT----- 1554  
 QY 543 GlyValIleGlyTYRThrHisSerSerValAspArgTYRAsnAlaIleSerAspLysIle 562  
 DB 1555 ---ACGTTTGGCTGGACACACTTAGTGTGATCCCTAAAAATCAATTTATACACTTTA 1611  
 QY 563 IleThrMetIleProAlaIleLysGlyAsnAsnLeuAspThrAsnSerLysValIleGln 582  
 DB 1612 ACTACCAAAATTCAGCTGTAAAGCGAATTCACCTGGAGCTCTTAAGCTTGTTCA 1671  
 QY 583 GlyProGlyHisThrGlyGlyAsnLeuValTYRLeuGlnSerGlnGlyArgLeuGlnIle 602  
 DB 1672 GGACCTGGTCATACAGAGGGAGTTTAAT-----GATTTCACAGATCATTTCAAAAT 1725  
 QY 603 ThrCysGluThrProAsnSerThrGlnSerTYRPhelIleArgLeuArgTYRAlaThrAsn 622  
 DB 1726 ACATGTCAACACTCAAAATTTTCAACATCTGATTTTATAGAAATTCGTTATGCTTCAAT 1785  
 QY 623 GlyAlaGlyAsnThrLeuProAsnIleSerLeuThrIleProGlyValIleGlyIlePro 642  
 DB 1786 GGAAGCGCAATACACGAGCTGTATTAATCTTAGATCCAGGGGTACACAGACTG--- 1842  
 QY 643 ProGlnArgLeuAsnAsnThrPheSerGlyTYRAsnTYRAsnAsnLeuGlnTYRGLYAsp 662  
 DB 1843 GGTATGGCACTCAACCCCACTTTTCTGCTACAGATTAATCAAAATTTAAATAAAGAT 1902  
 QY 663 PheGlyTYRPhelIlePheProSerThrValThrLeuProLeuAsnArgAsnIleProPhe 682  
 DB 1903 TTTCACTACTTAGAATTTCTAACAAGTGAATTTGCTTCAATCAAAATCAATCTCTT 1962  
 QY 683 IlePheAsnArgAlaAspVal---SerAsnSerIleLeuIleIleAspLysIleGluPhe 701  
 DB 1963 GGTGTTAATCGTTCGGATGTATATACAAACACACAGTACTTATGATTAATGATTT 2022  
 QY 702 IleProIleThrSerSerMetHisGlnAsnArgGlyLysGlnLysLeuGlnIleGln 721  
 DB 2023 CTGCCAATTAATCTGTTCTTAAGAGAGTAGAGAGAAACAAATAATTAAGAACAGTACAA 2082  
 QY 722 ThrLysIleAsnThrPhePheThrAsnHisThrLys 733  
 DB 2083 CAATTAATTAATCAATTTATGCAAAATCCTATATAA 2118  
 RESULT 6  
 ID AA081178 standard; DNA; 3543 BP.  
 XX

AC AA081178;  
 XX  
 DT 25-MAR-2003 (revised)  
 DT 12-AUG-1995 (first entry)  
 XX  
 DE B.t. toxin PS71M3 gene.  
 XX  
 KW Delta-endotoxin; crystal protein; biological control agent; Calliphoridae;  
 KW screw-worm; sheep blowfly; Lucilia; Phormia; Calliphora; insecticide;  
 KW pesticide; B.t; ss.  
 XX  
 OS Bacillus thuringiensis.  
 XX  
 PN MO9502694-A2.  
 XX  
 PD 26-JAN-1995.  
 XX  
 PF 13-JUL-1994; 94WO-US007902.  
 XX  
 PR 15-JUL-1993; 93US-00093199.  
 XX  
 PA (MYCO ) MYCOGEN CORP.  
 XX  
 PI Hickie LA, Payne J;  
 XX  
 DR MPI: 1995-067338/09.  
 DR P-PSDB; AAR63078.  
 XX  
 PS Disclosure, Page 36-38; 50pp; English.  
 XX  
 CC A library was constructed from Bacillus thuringiensis PS71M3 total  
 CC cellular DNA in lambda Gem-11. Plasmid pMYC1625, selected in Escherichia  
 CC coli, contained a 8.0 kb insert expressing a beta-endotoxin gene. This  
 CC was sequenced (AA081178). A cured, acrycallyiferous B.t. host carrying  
 CC pMYC1625 produced a 130 kDa crystal protein (AAR63078) toxic to  
 CC calliphorids. (Updated on 25-MAR-2003 to correct PN field.)  
 XX  
 SQ Sequence 3543 BP; 1278 A; 566 C; 612 G; 1087 T; 0 U; 0 Other;  
 Alignment Scores:  
 Pred. No.: 2,156-85 Length: 3543  
 Score: 1060.50 Matches: 271  
 Percent Similarity: 50.3% Conservative: 117  
 Best Local Similarity: 35.1% Mismatches: 279  
 Query Match: 27.4% Indels: 105  
 DB: 2 Gaps: 24  
 US-10-783-417-2 (1-735) x AA081178 (1-3543)  
 QY 1 MetAsnGlnAsnAsnAspAsnAsnGluTYRGLYIleIleAspSerHisThrSerProTYR 20  
 DB 1 ATGAATCTCTTATCAAAATTAATAATGAAATGAAACATTAAAGCTTCCACAAAAAATA 60  
 QY 21 PheProAsnArgAsnSerAsnAspSerArgTYRProTYRThrAsnAsnProAsnGlnPro 40  
 DB 61 -----AATATATCTTAATTAATATATACAGATATCCAAATAGAAATAGTCCAAACAAATTA 114  
 QY 41 LeuGlnAsnThrAsnTYRLeuGluTYRLeuAsnMetCysGlnGlyAsnThrGlnTYRGLY 60  
 DB 115 TTAACAAGTACAAATTAATTAAGATTGGCTCAATATAGTGCACAGAAATCAGCATGTGAT 174  
 QY 61 AsnAsnPheGluThrPheAlaSerAlaAspThrIleAlaValSerAlaGlyThrIle 80  
 DB 175 GGAAGATTTTGAACCTTTATGATAGT-----GGTGAACCTCAGTCCATATCTAT 225  
 QY 81 ValSerGlyTYRLeuLeuAlaGlyIleGlyGlyLeuThrSerIleSerGlyProIleGly 100  
 DB 226 GTAGTTGGGACCGGTACTGACTGTGTTCCGCTTCACACAC-----CCCTTAGGA 273  
 QY 101 IleIleGlyAlaIleIleIleSerPheGlyThrLeuIleThrValPheThrProAlaGly 120

```

Db      274 CTT-----GCTTATAGTGTGGTACATTAAATACAGTCTTTTCCAGCCCA 324
Qy      121 GluGlnAspLysThrValTrpThrGlnPheIleuMetGlyGluIlePheValAspThr 140
Db      325 GACCAATCTAACACA---TGGAGTCACTTTATACACAACTAAATAATTTATATAAAAA 381
Qy      141 ProLeuThrGlnSerIleuGlnIleuLysLeuGlnThrLeuGlnGlyPheArgGlnIle 160
Db      382 GAATAGAGATCAATATATATAAGTAAATGCTAATAAATTTTAAACAGGTCGTTAAATT 441
Qy      161 LeuGlnSerThrAsnThrAlaLeuAspAspTrpArgLysLeuLysArgLeuGlnAlaPro 180
Db      442 ATCAGACCTTATCATATATCATCCTTAAACATGGGAG-----AATATATCA 486
Qy      181 GlyLeuProProSerSerAlaLeuGlnGlnAlaAlaLeuThrLeuLysIleArgPheGlu 200
Db      487 AACCCACAAATATCTCAGAGATGTAAAGACACAAATCCAGTAACTTATACCATTTCA 546
Qy      201 AsnValHisAsnAspPheIleArgGluIlePro-----GlyPheGlnLeuGlnThrTyr 218
Db      547 AATGTCATTTCCAGAGCTGTAAACCTTGCTCTCTCTAATCCAGATTTGGCATTTACAT 606
Qy      219 LysThrLeuLeuLeuProIleTyrAlaGlnAlaAlaAsnPheHisLeuAsnLeuGln 238
Db      607 AACATACTAGTATATCTAGTATGTCACAGACGACAACTTACATCTGATTAATAAT 666
Qy      239 GlnGlyAlaGluLeuAlaAspGluTrpAsnIleAspIleHisProSerGlnIleGluPro 258
Db      667 CAAAGCCGTCAAATTTGAAAGCGTATTTAAAAACAAATCCACAAATTCGATTTATTAAGCT 726
Qy      259 AsnAlaGlyThrSerAspAspTyrTyrLysLeuLeuLysGluAsnIleProLysTyrSer 278
Db      727 ---TTGCCACACGCAATGATTATATTCAGATTTAGCATTAAGCATTAAGATTAACAT 783
Qy      279 AsnTyrGlyAlaAsnThrTyrArgThrGlyLeuLysAsnLeuAspGluPro----- 296
Db      784 AATATTTGTGTAACTATATAAAAAGATTTAAATTTAAATTAACACGCTCATAGT 843
Qy      297 -----AsnMetLysTrpSerIlePheAsnAspTyrArgArgTyrMetThrIle 312
Db      844 AATCTTGATGCAATATATACTGAACACATACAAATCGTATCGAACAATAATGCTACT 903
Qy      313 ThrValLeuAspThrIleSerGlnPheSerLeuTyrAspIleLysArgTyrArgAspSer 332
Db      904 GGTGATTTAGATCTGTTGCACTCTTCTCTAATTTAGTATGTAAGTAATAT----- 954
Qy      333 IleGlyGlyIleGluValLysGlyIleLysAsnGluLeuThrArgGlnIleTyrThrThr 352
Db      955 -----CCATAGGTGTCCAAATCTGAATCTGCAATCTGCAAAATTTATCAGGTA 999
Qy      353 GluIleAsnPheAspArgLeuProGlnLeuArgValGlnProAsnLeuAlaThrMetGlu 372
Db      1000 ---CTTAACCTTCGAAAGAAAGCCCTATATAATTTAT-----GACTTCCTCAATATCAAG 1050
Qy      373 TyrAsnLeuThrArgAlaSerPheLysLeuPheSerPheLeuGlnGlnPheIlePheTyr 392
Db      1051 GATTCACCTTACCGTAGACCG---CATTTATTTACTTGCTTGATTTCTTGAATTTTAT 1107
Qy      393 -----ThrGluAsnThrAsnPheGlyAsnArg----- 401
Db      1108 GAAAAAGCCAAACTACTCTTAATTTTTCACACACCATTAATATATATGTTTCATTAC 1167
Qy      402 ---LeuValGlyIleSerAsnArgAspAlaProThrTyrSerAsnThrIleThrGluThr 420
Db      1168 ACACCTTGATTAATATATCCAAATAATCTAGTGTGTTTGGAAATCACAATGATGATTA 1227
Qy      421 LeuTyrGlyGluArgThrGlySerProThrThrLysThrIleArgProPheGlnSerTyr 440
Db      1228 TTA-----AATCTCTGTTGGCACAATAATATTATTT----- 1263
Qy      441 LysValSerIleValThrAspArgLysSerProProValSerProIleGlnProHisPhe 460

```

```

Db      1264 -----TTT 1266
Qy      461 IleIleAsnGlnIleGluLeu-----TyrLeuAsnGlySerSerAsnAsnThr--- 476
Db      1267 TTATTAATATGTCATAGCTTATAGATTAATATATCTTAATGATTAATATATTTAGTAA 1326
Qy      477 -----LeuLysTyrSerAlaGly 483
Db      1327 ATGGAATTTTATTAATATAGTACTAGACTTTTGGAGAAAGAACTTAACAGAGACT 1386
Qy      484 SerLeuSerAsnTyrGln---AsnThrThrPhePheGlnPheProArgLysLysAspCys 502
Db      1387 GGGCAATTAAGCTTATGATGTAAATTAATAATTTTGGGTTTACCAATTTTAAACCAAGA 1446
Qy      503 AsnLeuValIleAspProGlyCysSerProAsnPheAsnAsnTyrSerHisIleLeuSer 522
Db      1447 GAGAAATCAAGAAATCCCTACCTTTTCCACATATGATTAAGTATATGATTAATTTATCA 1506
Qy      523 HisPheSerLeuPheThrTyrSerTyrValIleGlyLeuGlnIleLeuAspThr 542
Db      1507 TTTATTAAGAATCTTAGATATCCCTGCACACATATAAATCTCAAGGTAT----- 1554
Qy      543 GlyValLeuGlyTyrPheHisSerSerValAspArgTyrAsnAlaIleSerAspLysIle 562
Db      1555 ---ACGTTTGCTTGACACACTTAGTGTGATCTTAATAAATACAAATTTATACACATTTA 1611
Qy      563 IleThrMetIleProAlaIleLysGlyLysAsnLeuAspThrAsnSerLysValIleGlu 582
Db      1612 ACTACCCAAATTTCCAGCTGTAAAGCCAAATTCATCTGGAGACTGCTCTTAAGGTTTCCA 1671
Qy      583 GlyProGlyHisThrGlyLysAsnLeuValTyrLeuGlnSerGlnLysArgLeuGluIle 602
Db      1672 GGAAGCTGGTATACAGAGAGGAGATTAAAT-----GATTCAAAGATCATTTCAAAAT 1725
Qy      603 ThrCysGluThrProAsnSerThrGlnSerTyrPheIleArgLeuArgTyrAlaThrAsn 622
Db      1726 ACATGTCAACACTCAAAATTTTCAACAAATCGTATTTTAAGAAATTCGTTATGCTTCAAT 1785
Qy      623 GlyValGlyAsnThrLeuProAsnIleSerLeuThrIleProGlyValIleGlyIlePro 642
Db      1786 GAAAGCCCAATATACAGAGCTGTATTAATCTTACTATCCAGGGGTAGCAGAACTG--- 1842
Qy      643 ProGlnArgLeuAsnAsnThrPheSerGlyThrAsnTyrAsnAsnLeuGlnTyrGlyAsp 662
Db      1843 GGTATGAGCATCAACCCACACTTTTCTGTGACAGTTATAGAAATTTAAATATATAAGAT 1902
Qy      663 PheGlyTyrPheGlnPheProSerThrValThrLeuProLeuAsnArgAsnIleProPhe 682
Db      1903 TTTCAATCTTATAGAAATTTTCTTAACAGAGTGAATTTGCTCCAAATCAAAACATATCTCT 1962
Qy      683 IlePheAsnArgAlaAspVal---SerAsnSerIleLeuIleIleAspLysIleGluPhe 701
Db      1963 GGTGTTAATCGTTGGATGTATATACAAACACAAACAGTACTTATGATTAATAATTTGAAT 2022
Qy      702 IleProIleThrSerSerMetHisGlnAsnArgGlyLysGlnLysLeuGluThrIleGln 721
Db      2023 CTGCCAATTTATCTGTTCTATTAAGAGAGATTAAGAGAAACAAATTTGAAACAGTACAA 2082
Qy      722 ThrLysIleAsnThrPhePheThrAsnHisThrLys 733
Db      2083 CAAATTAATTAATACATTTATGCAAAATCTTATAAA 2118

```

RESULT 7  
 AAN93054  
 ID AAN93054 standard, DNA, 3940 BP.  
 AC AAN93054;  
 XX  
 AC 27-AUG-2003 (revised)  
 DT 25-MAR-2003 (revised)  
 DT 30-MAY-1990 (first entry)  
 XX  
 DE Delta-endotoxin crystal protein gene.

```

XX  Delta-endotoxin; crystal protein; insecticide; ss; PCC130;
KM  biological control agent.
XX
OS  Bacillus thuringiensis; israelensis.
FH  Key
FT  RBS      Location/Qualifiers
FT  RBS      879..884
FT  CDS      /label= Shine-Delgarno sequence
FT  CDS      891
FT  CDS      /*tag= a
FT  CDS      /product= "delta-endotoxin"
PN  BP296870-A.
XX
PD  28-DEC-1988.
XX
PF  24-JUN-1988; 88EP-00305772.
XX
PR  26-JUN-1987; 87US-00067653.
XX
PA  (DUPO ) DU PONT DE NEMOURS & CO E. I.
PI  Ellar DJ, Ward ES;
XX
PI  WPI; 1989-001332/01.
XX
DR  P-P5DB; AAP94035.
XX
PT  DNA fragment encoding insecticidal protein - obcd. from Bacillus
PT  thuringiensis sub species israelensis, and used in microorganisms and
XX  plant cells.
PS  Disclosure; Fig 5; 26pp; English.
XX
CC  The nucleotide sequence is an insert in plasmid PCC130. The delta
CC  endotoxin protein is insecticidal and can be used to control pests
CC  esp. mosquitoes. See also AAN93059. (Updated on 25-MAR-2003 to correct PA
CC  field.) (Updated on 27-ANG-2003 to correct OS field.)
SQ  Sequence 3940 BP; 1457 A; 603 C; 659 G; 1221 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:      3.06e-85      Length:      3940
Score:          1059.50      Matches:      271
Percent Similarity: 50.3%      Conservative: 117
Best Local Similarity: 35.1%      Mismatches:  279
Query Match:      27.4%      Indels:      105
DB:              1          Gaps:      24

US-10-783-417-2 (1-735) x AAN93054 (1-3940)
QY  1 MetAsnGlnAsnAspAsnAsnGluTyrGluIleIleAspSerHisThrSerProTyr 20
DB  891 ATGAATCCCTTAACAAATTAATAAATGAATATGAATACATTAATGCTTCACAAAAAAATTA 950
QY  21 PheProAsnArgAsnSerAsnAspSerArgTyrProTyrThrAsnAsnProAsnGlnPro 40
DB  951 -----AATATATCTAATAATATTATACAGATATCCATAGAAATAGTCCAAATAACAATTA 1004
QY  41 LeuGlnAsnThrAsnTyrIleGluTyrIleAsnMetCysGlnGluIleAsnThrGlnTyrGly 60
DB  1005 TTACAAAGTACAAATTAATAAAGATGCTCAATATATGTCACAGATCAGCATGATGGT 1064
QY  61 AspAsnPheGluThrPheAlaSerAlaAspThrIleAlaAlaValSerAlaGlyThrIle 80
DB  1065 GCGAATTTTGAACCTTTATATGATAGT-----GGTGAATCTCAGTGCCTATATATT 1115
QY  81 ValSerGlyThrLeuLeuAlaGlyIleGlyIleuThrSerIleSerGlyProIleGly 100
DB  1116 GTATGTGGGACCGCTACTGACTGCTTGGGTTCCAGACACA-----CCCTTAAGGA 1163
QY  101 IleIleGlyAlaIleIleIleSerPheGlyThrLeuIleThrValPheTrpProAlaGly 120

```

```

DB  1164 CTT-----GCTTAATAGGTTTGGTACATTAATACAGTTCTTTTCAGCCCAA 1214
QY  121 GluGlnAspLysThrValTrpThrGlnPheIleIleIleMetGlyGluIlePheValAspThr 140
DB  1215 GACCAATCTMACACA---TGAGTGACTTATTAACAACAACTAAATATTATTAACAAAA 1271
QY  141 ProLeuThrGluSerIleIleGluIleuLysGluIleThrLeuGluIlePheArgGlnIle 160
DB  1272 GAATATGACATCAACATATATTAAGTAATGCTAATTAATTTAAACAGGTCGTTAATGTT 1331
QY  161 LeuGlnSerTyrAsnThrAlaLeuAspAspThrPArgLysLeuLysArgLeuGlnAlaPro 180
DB  1332 ATCAGACCTTATCATATATACCTTAAACATGGAG-----AATATCCA 1376
QY  181 GlyLeuProProSerSerAlaLeuGlnGlnAlaAlaLeuThrLeuLysIleArgPheGlu 200
DB  1377 AACCCACAAAATATCATCAGATGTAAAGACACAAATCCAGCTAGTTCATTACCATTTTCAA 1436
QY  201 AsnValHisAsnAspPheIleArgGluIlePro-----GlyPheGlnLeuGluThrTyr 218
DB  1437 AATGTCATTCACAGCTGTAAACTTGTCTCTTAATCTTAAGTATTCGATTCATTAATA 1496
QY  219 LysThrLeuLeuLeuProIleTyrAlaGlnAlaAlaAsnPheHisLeuAsnLeuGln 238
DB  1497 AACATATAGTATTTATCTAGTATATGACAGACAGCAAACTTACATCTGACTATTAAT 1556
QY  239 GlnGlyAlaGluLeuAlaAspLysThrAsnAlaAspIleHisProSerGlnIleGluPro 258
DB  1557 CAGCCCGTCAAAATTTGAACCGTATTTAAACAAATCGAACAAATTCGATTTATTAAGACCT 1616
QY  259 AsnAlaGlyThrSerAspAspTyrTyrLysLeuLeuLysGluAsnIleProLysTyrSer 278
DB  1617 ---TTGCCAACAGCATTTGATTAATTAATCCAGATTAAGCTTAAGAAAGATTACACT 1673
QY  279 AsnTyrCysAlaAsnThrTyrArgThrGlyLeuLysAsnLeuArgAspGluPro----- 296
DB  1674 AATTATGTGTAAACAACCTTAATAAAGATTAATTAATTAATAAAGACGCTGATAGT 1733
QY  297 -----AsnMetLysTrpSerIlePheAsnAspTyrArgArgTyrMetThrIle 312
DB  1734 AATCTTGATGAAATATTAACCTGAACACATCAATCAATCAATCAACAAATATGACTACT 1793
QY  313 ThrValLeuAspThrIleSerGlnPheSerLeuTyrAspIleLysArgTyrArgAspSer 332
DB  1794 GCTGATTAAGATGTTGTGCACTCTTCTTAATTTATGATGATGATTAATAT----- 1844
QY  333 IleGlyGlyIleGluValLysGlyIleLysAsnGluLeuThrArgGluIleTyrThrThr 352
DB  1845 -----CCAAATAGGTGTCCAAATCTGAATCTGAACTGAGAAATTTATTCAGCTA 1889
QY  353 GluIleAsnPheAspArgLeuProGlnLeuArgValGlnProAsnLeuAlaThrMetGlu 372
DB  1890 ---CTTAATCTTGAAGAAAGCCCTTAATAATTTAT-----GACTTCAATATCAAGAG 1940
QY  373 TyrAsnLeuThrArgLysPheLysPheLysPheSerPheLeuGlnPheIlePheTyr 392
DB  1941 GATTCATCTTACAGTGAACCG---CATTTATTTACTGCTGATTCCTTGAATTTTAT 1997
QY  393 -----ThcGluAsnThrAsnPheGlyAsnArg----- 401
DB  1998 GAAAAAGCGCAACATCTCTTAATATTTTTCACACGCAATTAATATGTTTCATTAC 2057
QY  402 ---LeuValGlyIleSerAsnArgAspAlaProThrTyrSerAsnThrIleThrGluThr 420
DB  2058 ACACCTGATTAATATATCCCAAAATCTAGTGTGTTTGGAAACACAACTAATTAATAA 2117
QY  421 LeuTyrGlyLysArgTyrSerProThrLysThrIleArgProPheGluSerTyr 440
DB  2118 TTA-----AATTCCTTGTGTTGCAACAAATATTATTAAT----- 2153
QY  441 LysValSerIleValThrAspArgLysSerProValSerProIleGlnProHisPhe 460

```

Db	2154	----	TTT	2156
Qy	461	llelleangnlleluLeu-----TyrleuangelSerSerAenAenThr--	476	
Db	2157	TTATTAATGCTAAGCTTAGATATATATAATCTAAATGATTTATTAATTAATGATAAA	2216	
Qy	477	-----LeuLysTyrSerAlaGlYGlY	483	
Db	2217	ATGATTTTTTTATTAACTAATGTAAGTACCTTTGGAGAAAGAACTTACGCGAGATCT	2276	
Qy	484	SeiruseAenTYrGln---AsnThrThrphepneglInphepOArgLYLysAspCys	502	
Db	2277	GCGCAATTAATCTAATGATGATAATATAAATAATTTCCGGTTACCAATCTTAAACGAGA	2336	
Qy	503	AenLeuVallleapProGlyCysSerProAenpheAenAenTYrSerHislleLeuSer	522	
Db	2337	GAGATCAAGGAACCCCTACCTTTTTCCAACATATGATTAATCTATAGCATATTTTATCA	2396	
Qy	523	HisPheSerleuAenpHeThrTYrSerTYrVallleGlyLeuGlnleuGlnlleuAenpThr	542	
Db	2397	TTTATTAATAAGCTTTAGTATCCCTGGCAACATATATAAACTCAAGGTAT-----	2444	
Qy	543	GlyValleuGlyTYrTpThrHisSerSerValAspArgTYrAenAlaHisSerAspLYsile	562	
Db	2445	---ACGTTGCTGTGACACACTCTAGTGTGATCTCTAAAAATACATTTATACACATTA	2501	
Qy	563	lleThrMetlleProAlalleGlyGlyAenAenleuAenpThrAenSerLYsVallleGlu	582	
Db	2502	ACTACCCAAATTCACGCTGTAAAAAGCGAATTCACCTGGACCTGGCTTCAAGCTTTGTTCAA	2561	
Qy	583	GlyProGlyHisThrGlyLYsAenLeuValTYrleuGlnSerGlnGlyAenleuGlnlle	602	
Db	2562	GGACTGTGTATACAGGAGGGGATTTAATT-----GATTTCAAAAGATCATTTCAAAATT	2615	
Qy	603	ThrCysGlnThrProAenSerThrglnSerTYrPheHlleArgLeuArgTYrAlaThrAen	622	
Db	2616	ACATGTCAACACTCAAAATTTTTCACAAATCGATTTTATTAAGAAATTCGTATAGCTTCAAAT	2675	
Qy	623	GlyAlaGlyAenThrleuProAenlleSerleuThrilleProGlyVallleGlyIlePro	642	
Db	2676	GGAACCGGAATATCTCGAGCTGTATTAATCTTAGATCCAGGGGTAGCGAAGACTG---	2732	
Qy	643	ProGlnAenleuAenAenThrPheSerGlyThyrAenTYrAenAenleuGlnTYrGlyAer	662	
Db	2733	GGTATGGACTCAACCCCACTTTTCTGTACAGATTATACGAATTTTAAATATTAATAAT	2792	
Qy	663	PheGlyTYrPheGlnPheProSerSerThrValThrleuProLeuAenAenAenlleProPhe	682	
Db	2793	TTTCAGTACTTAAGATTTTCTTAACGAGGTGAATTTGCTCCAAATCAAAACATATCTCTT	2852	
Qy	683	llePheAenAenAlaAspVal---SerAenSerilleuIlelleAspLYsilleGluPhe	701	
Db	2853	GGTTTATTCGTTCGAGTGTATATACAAACACACACAGTACTTATTTGATTAATTTGAATTT	2912	
Qy	702	lleProIleThrSerSerMetHisGlnAenAenArgGlnLYsGlnLYsleuGlnThrilleGln	721	
Db	2913	CTGCCAATTAATCTCGTTTATTAAGAGAGAGATGAGAGAAACAAATTTGAAACAGTACAA	2972	
Qy	722	ThrLYsilleAenThrPhePheThrAenHisThrLYs	733	
Db	2973	CAAAATTAATTAATACATTTTATGCAAAATCTATATAAA	3008	
RESULT 8				
AAN81490				
ID	AAN81490 standard; DNA; 4934 BP.			
XX	AAN81490;			
AC				
DT	24-OCT-2003 (revised)			
DT	05-NOV-1990 (first entry)			
XX				
DE	Insecticidal (Diptera) protein gene.			

KW	Insecticidal protein gene; Diptera; Aedes; ss.
XX	
OS	Bacillus thuringiensis serovar israelensis.
XX	
FH	Key Location/Qualifiers
FT	CDS 1..4908
FT	/tag= a
FT	/product= "insecticidal protein"
FT	620..1368
FT	intron
FT	/tag= b
XX	
PV	JP63230090-A.
XX	
PD	26-SEP-1988.
XX	
PF	19-MAR-1987; 87JP-00066844.
XX	
PR	19-MAR-1987; 87JP-00066844.
XX	
PA	(SUMO ) SUMITOMO CHEM IND KK.
XX	
DR	WPI, 1988-311968/44.
DR	P-PESDB; AAP81034, AAP82589.
XX	
PT	New insecticidal protein of bacillus thuringiensis israelensis strain -
PT	prepd. by isolating plasmid contg. insecticidal protein gene by forming
XX	gene library from plasmid deoxyribonucleic acid, etc.
XX	
PS	Disclosure; Page ?; 9pp; Japanese.
XX	
CC	A plasmid contg. the gene is isolated by forming a gene library from
CC	plasmid DNA of B. thuringiensis israelensis (HD 522 strain, USA Goldberg
CC	ON60) followed by screening with anti-israelensis insecticidal protein
CC	IG. The insecticidal protein is highly effective against Diptera, esp.
CC	Aedes. (Updated on 24-OCT-2003 to standardise OS field)
XX	
SQ	Sequence 4934 BP; 1797 A; 749 C; 877 G; 1511 T; 0 U; 0 Other;
Alignment Scores:	
Pred. No.:	4,196-85 Length: 4934
Score:	1059.50 Matches: 271
Percent Similarity:	50.4% Conservative: 118
Best Local Similarity:	35.1% Mismatches: 278
Query Match:	27.4% Indels: 105
DB:	1 Gaps: 24
US-10-783-417-2 (1-735) x AANB1490 (1-4934)	
OY	1 MetcAenGlnaAsnaAspaNsangLutYrgUlleIeAspSerHisTherSeProTYr 20
Db	1369 ATGAATCCTTATCAAAATAAAAATTAATGAATTAAGTTACGCACAAAAAATTAA 142
OY	21 PhePrOAAnaAgAsnSerAsnaAspSerXgrTYrProTYrThrAsnaAsProkAngInPro 40
Db	1429 -----AATATATCTCAATTAATTAATTCAGAATGCCAATGAAATAGCAAAACAATTAA 148
OY	41 LeuglNaEntThrAsnTYrLeuglyTrleuAuEnMetCySGlngLYaSnThrgInTYrgLY 60
Db	1483 TTACAAAGTCAAAATTAATTAAGATTGGCTCAATATGTGTCACACAAATACGCGATGTGGT 154
OY	61 AspaEnPhegiuthrPhelaSeraAlaaspThrilealaalavaSeraNlaaglythrile 80
Db	1543 GGAGATTTTGAAACTTTTATGTATGATGT-----GGTGAACCTCAGTCCCTATATATT 159
OY	81 ValSerGIyThrLeuEuaNlaGIyTLegLyGLeuThrSerTleSerGIyProIIegLy 100
Db	1594 GTAGTGGGACCGTACTGCTGGTTCCGGTTCACACA-----CCCTTAGGA 164
OY	101 IleleIeGIyAlaIleIleIleSerPheGIyThrleuUIleThrValPheTrProNlaGIy 120
Db	1642 CT-----GCTTTAATACGTTTTGGTACATTAAATCCAGTCTTTTCCAGCCCAA 169
OY	121 GluGlnaPlyeThrValIleThrThrgInPheIleIysMeTGlYglUllePheValaSPThr 140







QY 503 AaenLeuVal11eAspProGlyCysSerProAenPheAaenVal1SerHis1leuSer 522  
 Db 1476 GAAACCAAGCAAGCAACCGTGTCCGACCTACGACATACGACCAATCCGAGC 1535  
 QY 523 HisPheSerLeuPheThrTySerTyVal11leGlyLeuGln1leuAspThr 542  
 Db 1536 TTCATCAAAAGCCTGACATCCCGGCAACCTACAAAACCAAGTGATAC----- 1583  
 QY 543 GlyVal1leuGlyTyPheHisSerSerValAaPargTyAsnAla1leSerAplySile 562  
 Db 1584 ---ACCTTCGCGTGAACCAAGCAAGCGTGCAGCCGAAACCAATCTACACCCACCTG 1640  
 QY 563 1leThrMet1leProAla1leLeuGlyAaenAaenLeuAspThrAaenSerTyVal11leGln 582  
 Db 1641 ACCACCCAAATCCCGCGGTGAAGCAAGCAAGCTTGCGACGCAAGAAAGTGCCAA 1700  
 QY 583 GlyProGlyHis1ThrGlyGlyAaenLeuVal1TyPheGlnSerGlnGlyAArgLeuGln 602  
 Db 1701 GGTCCGGGTCAACCGCGGTGACCTGATC-----GACTTCAAGACCACTTCAAAATC 1754  
 QY 603 ThrCysGluThrProAaenSerThrGlnSerTyPhe1leAArgLeuAArgTyAlaThrAaen 622  
 Db 1755 ACCTCACACACAGCAACTTCCACAAAGCTACTTCATCCGATCCGTTACCGGAGCAAC 1814  
 QY 623 GlyAlaGlyAaenThrLeuProAaen1leSerLeuThr1leProGlyVal11leGly1lePro 642  
 Db 1815 GCGAGCGCAACACCGCGCGGTGATCAACCTGAGATCTCTGCGTGGCAGAACCTG--- 1871  
 QY 643 ProGlnAArgLeuAaenThrPheSerGlyThrAaenTyAaenAaenLeuGlnTyGlyAaP 662  
 Db 1872 GGTATGCACTGAACCGGACCTTCAAGCGGTACCGACTACACCACTGAATATCAAAAGAC 1931  
 QY 663 PheGlyTyPheGlnPheProSerThrVal1ThrLeuProLeuAaenAArgAaen1leProPhe 682  
 Db 1932 TTCATACCTGAGATTGAGCAACGAAGTGAATTCGCGCCGACCAACCAATCAAGCTG 1991  
 QY 683 1lePheAaenAArgAlaAspVal1---SerAaenSer1leLeu1le1eAspTySileGluPhe 701  
 Db 1992 GGTTCACACCTGACGACCTGTACACCAACACACCGTCTGTATGCAAAATCCAGTTTC 2051  
 QY 702 1lePro1leThrSerSerMetHisGlnAaenAArgGlyGlyGlnTySileGlnTyThr1leGln 721  
 Db 2052 CTGCGGATCACCGTACGATCCGCGAAGACCGTGAAGAAACAAACTGGAACCGTGCA 2111  
 RESULT 10  
 ADP71295 standard; DNA; 2246 BP.  
 XX ADP71295;  
 AC ADP71295;  
 XX  
 DT 09-SEP-2004 (first entry)  
 XX  
 DE Codon-modified Cry4A gene #2.  
 XX  
 KM codon optimisation; codon-modified; Cry4A; gene; ds.  
 XX  
 OS Synthetic.  
 OS Unidentified.  
 XX  
 PN JP2004166574-A.  
 XX  
 PD 17-JUN-2004.  
 XX  
 PF 19-NOV-2002; 2002JP-00335440.  
 XX  
 PR 19-NOV-2002; 2002JP-00335440.  
 XX  
 PA (UYOK-) UNIV OKAYAMA.  
 XX  
 DR WPI; 2004-445549/42.  
 XX  
 PT Novel codon-modified gene encoding Cry4A protein, the modification being substitution with most frequently used codons in host microorganisms such

PT that resulting gene has high translation rate in the microorganism.  
 XX  
 PS Disclosure; SEQ ID NO 2; 20bp; Japanese.  
 XX  
 CC The invention comprises a codon-modified gene which codes for Cry4A  
 CC protein, the gene is modified by substitution with the most frequently  
 CC used codon in the host microorganism, so that the modified gene will have  
 CC a high translation rate in the host microorganism. The codon-modified  
 CC gene of the invention is useful for the production of Cry4A protein in  
 CC large quantities. The present DNA sequence represents a codon-modified  
 CC Cry4A gene of the invention.  
 XX  
 SQ Sequence 2246 BP; 711 A; 728 C; 457 G; 350 T; 0 U; 0 Other;  
 Alignment Scores:  
 Pred. No.: 3,928-83 Length: 2246  
 Score: 1032.50 Matches: 265  
 Percent Similarity: 50.34 Conservative: 117  
 Best Local Similarity: 34.94 Mismatches: 273  
 Query Match: 26.74 Indels: 105  
 DB: 12 Gaps: 24  
 US-10-783-417-2 (1-735) x ADP71295 (1-2246)  
 QY 1 MetAaenGlnAaenAaenAaenGluTyGlyLeu1leAaenSerHis1ThrSerProTy 20  
 Db 146 ATGAACCCGCTACCAAAACAAAGCAATACGAACCCGTAACGCCAGCAAGAAACCTG 205  
 QY 21 PheProAaenAArgAaenSerAaenAaenAaenAaenAaenAaenAaenAaenAaenAaen 40  
 Db 206 -----AATCATGACAAACATACACCGCTTACCGATCCGATGCAAAACAGCCCGAACAACCTG 259  
 QY 41 LeuGlnAaenThrAaenTyTyGlyLeuThrLeuAaenMetCysGlnGlyAaenThrGlnTyGly 60  
 Db 260 CTGCAAAACACCAACCAATCAAAAGCAAGCTGTAACAGTGTCCCAACAAACAAATACCGC 319  
 QY 61 AaenPheGluThrPheAlaSerAlaAaenThr1leAlaAaenValSerAlaGlyThr1le 80  
 Db 320 GCGACCTTCAAACTTATCATGACAGC-----GGTGAACTAAGCCGCTACACATC 370  
 QY 81 ValSerGlyTyThrLeuAaenAaenGlyTyGlyLeuThrSer1leSerTyPro1leGly 100  
 Db 371 GTGTCGCGACCGTCTGACCGGTTCGCTTCAACAC-----CCGCTGGGC 418  
 QY 101 1le1leGlyAla1le1le1leSerPheGlyTyThrLeu1leThrValPheThrProAlaGly 120  
 Db 419 CTG-----GCGCTGATCGTTCGTTCCGTACCCGTCGCTCTGTCCTGCGCGCA 469  
 QY 121 GlnGlnAaenTyThrVal1ThrThrGlnPhe1leTyMetGlyGln1lePheValAaenThr 140  
 Db 470 GACCAAGCAACACCT--TGGAGCACTTATCATCCCAACCAACCAACATTCATCAAAAAA 526  
 QY 141 ProLeuThrGlnSer1leTyGlnLeuTySileGlnThrLeuGlnGlyPheAArgGln1le 160  
 Db 527 GAAATCGAAGACCTTATCATGACGAAGCAAGCAAAATCTGTAACCGTACCTCAACGTG 586  
 QY 161 LeuGlnSerTyAaenThrAlaLeuAaenAaenAaenAaenAaenAaenAaenAaenAaen 180  
 Db 587 ATCAGCACTTACCAACACCACTGAAACCTTGGA-----AACCAACCG 631  
 QY 181 GlyLeuProProSerSerAlaLeuGlnGlnAla1leAaenThrLeuTySileAArgPheGlu 200  
 Db 632 AACCCGCAAAACCAAGACGTGCTGATCCCAATTCACCTGCTGACCTACCACTTCCAA 691  
 QY 201 AaenValHisAaenAaenPhe1leAArgGlu1lePro-----GlyPheGlnLeuGluTyThrTy 218  
 Db 692 AACGTCATCCCGGAACGTGTGAACGTCGCCGGAACCGGAGCACTGCGACATACATAC 751  
 QY 219 LysThrLeuLeuLeuPro1leTyAlaGlnAlaAaenPheHisLeuAaenLeuGln 238  
 Db 752 AACATCTGTGTCTGACGAGCTACCGCAACAGGAACCTGCACTTCAACGTCGTGAC 811  
 QY 239 GlnGlyAlaGluLeuAlaAaenAaenAaenAaenAaenAaenAaenAaenAaenAaenAaen 258

```

Db      812 CAAGGGCTCAATTGAGGCTTACTGTAACAAACCCGCAATTCGACTCCTCAGCCG 871
      259 AsnAlaGlyThrSerAspAspTyrTyrIleuLeuLysGluAsnIleProLysTyrSer 278
      872 ---CTGCCGACCGCATGACTACTACCCGGTGTGACCAAAAGCAATGAGACTACACC 928
      279 AsnTyrGlyAlaAsnThrTyrArgThrGlyLeuLysAsnLeuArgAspGluPro----- 296
      929 AACTACTGGGTGACCACTCAACAAAAAGCCCTGAACCTGATCAAAACCAACCCCGACAGC 988
      297 -----AsnMetLysTyrPheSerIlePheAsnAspTyrTyrArgTyrMetThrIle 312
      989 AACTGACGCGCAACATCAACTGAAACACCTACCAACACCTACCGCAACCAAAATGACCAAC 1048
      313 ThrValLeuAspThrIleSerGlnPheSerLeuTyrAspIleLysArgTyrArgAspSer 332
      1049 GCGGTGTGACCTGGTGGCTGTTCCTCCGAACTACGACGTGCGTAAATAC----- 1099
      333 IleGlyGlyIleGluValLysGlyIleLysAsnGluLeuThrArgGluIleTyrThrThr 352
      1100 -----CCGATCGGTGTCCAAAGCGAAGTGAACCCGGGAATCTACCAAGTC 1144
      353 GluIleAsnPheAspArgLeuProGlnLeuArgValGlnProAsnLeuAlaThrMetGlu 372
      1145 ---CTGAACCTTCGAAAGAAAGCCCGTACAAATACTAC-----GACTTCCAAATACCAAGAA 1195
      373 TyrAsnLeuThrArgAlaSerPheLysLeuPheSerPheLeuGlnGlnPheIlePheTyr 392
      1196 GACAGCTGACCCGTGCGCCG---CACTGTTCACCTGCTGAGACGCTGAACTTCTAC 1252
      393 -----ThrGluAsnThrAsnPheGlyAsnArg----- 401
      1253 GAAAAAGCGCAACACCCCGAACAACCTTCTCACAGCACTACACATGTTCCACTAC 1312
      402 ---LeuValGlyIleSerAsnArgAspAlaProThrTyrSerAsnThrIleThrGluThr 420
      1313 ACCCTGGAACAACATCAGCCCAAAAGACGCGTGTGCGACACCAAGCGTACCCACAAA 1372
      421 LeuTyrGlyLysArgThrGlySerProThrThrLysThrIleArgProPheGluSerTyr 440
      1373 CTG-----AAAGCTGGGTCTGGCAACCAACATCTACATC----- 1408
      441 LysValSerIleValThrAspArgLysSerProValSerProIleGlnProHisPhe 460
      1409 -----TTC 1411
      461 IleIleAsnGlnIleGluLeu-----TyrLeuAsnGlySerSerAsnAsnThr--- 476
      1412 CTGCTGAACGTTCATCAGCTGACCAAAATCTGAAACGACTACACACATCTCAGCAAA 1471
      477 -----LeuLysTyrSerIleGly 483
      1472 ATGCACTTCTTCATACCAACGCTACCCGCTCTGAAAAAAGACGACGACGAGC 1531
      484 SerLeuSerAsnTyrGln---AsnThrThrPhePheGlnPheProArgLysLysAspCys 502
      1532 GCGCAAAATCACTAGAGGTGAACAAAAACATCTTCGGCTGCGCATCTGAAAGTGGC 1591
      503 AsnLeuValIleAspProGlyCysSerProAsnPheAsnAsnTyrSerIleIleLeuSer 522
      1592 GAAAAACCAAGCAACCCGACCTGTCCGACCTACGACCACTACAGCACAATCTCGAGC 1651
      523 HisPheSerLeuPheThrTyrSerTyrValIleGlyLeuGlnIleLeuAspThr 542
      1652 TTCATCAAAAGCCTGACATCCCGCAACCTACAAACCCCAAGTATAC----- 1699
      543 GlyValLeuGlyTyrThrHisSerSerValAspArgTyrAsnAlaIleSerAspLysIle 562
      1700 ---ACCTTGGGTGACCAACGACGCGTCCGACCGGAAAAACACATCTACACCCACTG 1756
      563 IleThrMetIleProAlaIleLysGlyAsnAsnLeuAspThrAsnSerIysValIleGlu 582

```

```

Db      1757 ACCACCCAAATCCCGCGCGTGAAGCAAGACCTGGCGACCGGACGAGCAAAAGTGTCCA 1816
      583 GlyProGlyHisThrGlyGlyAsnLeuValTyrLeuGlnSerGlnGlyArgLeuGluIle 602
      1817 GGTCCGGTTCACCCCGCGGTGACTATC-----GACTTCAAAAGACCACTTCAAAATC 1870
      603 ThrCysGluThrProAsnSerThrGlnSerTyrPheIleArgLeuArgTyrAlaThrAsn 622
      1871 AACTGCCAACACAGCAACTTCCAAACAAAGACTTCACTCCGATCCGATCCGTTACGCGAGCAAC 1930
      623 GlyAlaGlyAsnThrLeuProAsnIleSerLeuThrIleProGlyValIleGlyLysPro 642
      1931 GCGACCGCAACACCCCGCGGTGATCACTGAGCATCCCTGGGTGCGAGAACTG--- 1987
      643 ProGlnArgLeuAsnAsnThrPheSerGlyThrAsnTyrAsnAsnLeuGlnTyrGlyAsp 662
      1988 GGTATGGACTGAAACCCGACCTTCAGCGGTACCGACTACACCACTGAATACAAAGAC 2047
      663 PheGlyTyrPheGlnPheProSerThrValThrLeuProLeuAsnArgAsnIleProPhe 682
      2048 TTCCAATACCTGAGATTACAGCAACGAGTGAATTCGCGCGCAACCAAAACATCAGCCTG 2107
      683 IlePheAsnArgAlaAspVal---SerAsnSerIleLeuIleIleAspLysIleGluPhe 701
      2108 GTGTCAACCGGTAGCGAGGTGTACACCAACCAACCGGTGATCGACAAATCGAGTTC 2167
      702 IleProIleThrSerSerMetHisGlnAsnArgGlyLysGlnLysLeuGluThrIleGln 721
      2168 CTGCCGATCACCCGTGACATCCGCAAGACCTGTAAAAACAAAACTGGAACCGTGCAA 2227

RESULT 11
AA014670
ID   AA014670 standard; DNA; 2061 BP.
XX
XX
AC   AA014670;
XX
DT   27-AUG-2003 (revised)
DT   25-MAR-2003 (revised)
DT   04-FEB-1992 (first entry)
XX
DE   Dipteran active toxin gene.
XX
KM   Insecticide; B.c; crystal; delta endotoxin; cryIIC; ss.
OS   Bacillus thuringiensis serovar morrisoni.
XX
XX
FH   Key
FT   CDS
FT   1..2028
FT   /*tag= a
XX
XX
PE   EP457498-A.
XX
PD   21-NOV-1991.
XX
PF   09-MAY-1991; 91EP-00304180.
XX
PR   15-MAY-1990; 90US-00524255.
PR   01-OCT-1990; 90US-00590903.
XX
PA   (MYCO ) MYCOGEN CORP.
XX
PI   Sick AJ;
XX
DR   WPI, 1991-341902/47.
DR   P-P8DB; AAR14374.
XX
PT   Bacillus thuringiensis genes encoding diptera-active toxins - and
XX   transformed microbes used to control insects in various environments.
XX   Claim 1; Page 15; 20pp; English.
XX
CC   The sequence was obtd. from plasmid pMYC1636 which was isolated from a
XX   genomic library prepd. from DNA from B.c. PS71M3 [from B.c. PS71M3-69

```

CC (NRRL B-18515)]. It is related to the cryIYC from B.c. var. israeliensis.  
 CC The gene encodes a 77 kD protein. Microorganisms transformed with the DNA  
 CC may be administered to dipteran insects or their environments, the  
 CC expressed toxins acting as an insecticide. See also AA014669-Q14672.  
 CC (Updated on 25-MAR-2003 to correct PA field.) (Updated on 27-AUG-2003 to  
 CC correct OS field.)

XX Sequence 2061 BP; 779 A; 331 C; 319 G; 632 T; 0 U; 0 Other;

# Alignment Scores:

Pred. No.:	Length:	Score:	Matches:	Conservative:	Mismatches:	Best Local Similarity:	Query Match:
3,866-71	2061	899.50	247	126	270	32.0%	23.2%
							Gaps: 30

US-10-783-417-2 (1-735) x AA014670 (1-2061)

```

QY 1 MetAenGlnAsnAsnAsnAsnAsnAsnGluTyrGluIleIleAspSerHisThrSerProTyr 20
Db 1 ATGAATCCATATCAAAATATGAATATGAATATGAATATTCATATGCTCATCCATGAGTTT 60
QY 21 PheProAsnArgAsnSerAsnAsnSerArgTyrProTyrThrAsnAsnProAsnGlnPro 40
Db 61 -----AGCAAGTCTAATATCTATCTATGATATCCATAGCAATATAGCAAAATACACCA 114
QY 41 LeuGlnAsnThrAsnTyrLysGluTyrPheLeuAsnMetCysGlnGlyAsnThrGlnTyrGly 60
Db 115 CTGAAATACACGAAATATCAAAAGATTGGCTCAATGTGTGTCAGATATATCAACAATATGCG 174
QY 61 AspAsnPheGluThrPheAlaSerAlaAspThrIleAlaIleValSerIleGlyThrIle 80
Db 175 AATATATGCGGGAATTTTGTAGTCTGAACATATTTGGAAGTATGTCAGATATATAT 234
QY 81 ValSerGlyThrLeuLeuAlaGlyIleGlyGlyLeuThrSerIleSerGlyProIleGly 100
Db 235 GTAGTAGAGAACTATGTAGGA-----GCTTTTGCTGCCCT----- 270
QY 101 IleIleGlyAlaIleIleIleSerPheGlyThrLeuIleThrValPheThrProAlaGly 120
Db 271 GCTTAGAGTGCAGATATATATCTTTTGGGACCTTTGTCCTCCATCTTTTGG--CAAGGA 327
QY 121 GluGlnAspLysThrValThrPheGlnPheIleLysMetGlyGluIlePheValAspThr 140
Db 328 TCTGACCTGCAAAATGTTTGGCAGATTTGTTAAACATCGA-----GGAAGG 375
QY 141 ProLeuThrGlu-----SerIleLysGlnLeuLysLeuGlnThrLeuGluGlyPhe 157
Db 376 CCTATACAAAGAAATATGATATAAAACATATATATGTAATCTTCTATCTGTAACACTATA 435
QY 158 ArgGlnIleLeuGlnSerTyrAsnThrAlaLeuAspAspTyrArgLysLeuLysArgLeu 177
Db 436 AAAAATCAACCTTGATTAATATCAAGAAATTTTTCGATTAATATGGAGCAGACAGCT----- 489
QY 178 GlnAlaProGlyLeuProProSerSerAlaLeuGlnGlnAlaAlaLeuThrLeuLysIle 197
Db 490 -----ACACACGCTAATGCTTAAGCAGTATGATCATGATCTCTTTACT 528
QY 198 ArgPheGluAsnValHisAsn-----AspPheIleArgGluIleProGlyPheGln 214
Db 529 ACCCTTAGAAGCTTATATGATTAAGATTATGATATGTTAAAAATATATGCTAGCTATGCA 588
QY 215 LeuGluThrTyrLysThrLeuLeuLeuProIleTyrAlaGlnAlaAlaAsnPheHisLeu 234
Db 589 ATACCAACA-----CTCCGCGATATGCAACAATATGCTAGCTAGCTAGCTAGCTAGCT 633
QY 235 AsnLeuLeuGlnGlnGlyAlaGluLeuAlaAspGluThrPheAsnAlaAspIleHisProSer 254
Db 634 AATTATTTAAACATGCTGCTACTTATCAATATATG-----CTGCAAAATCA 684
QY 255 GlnIleGluProAsnAlaGlyThrSerAspAspTyrTyrLys-----LeuLeuLysGluAsn 273

```

```

Db 685 GGTATTAATCCAAAGTACTTTCATTCATCTATATCTATACAGGCTATTTAAACGTAA 744
QY 274 IleProLysTyrSerAsnTyrCysAlaAsnThrTyrArgThrGlyLeuLysAsnLeuArg 293
Db 745 ATACAAAGATATATGACTATATGATTTATACAAAGTACATGACGAGCTATATCTATGATTA 804
QY 294 AspGluProAsnMetLysIlePheSerIlePheAsnAspTyrArgArgTyrMetThrIleThr 313
Db 805 ACTATATACAAAGCAATGGAATATGATATATATATCTATACCGTTTGAATGACTCTAACT 864
QY 314 ValLeuAspThrIleSerGlnPheSerLeuTyrAspIleLysArgTyrArgAspSerIle 333
Db 865 GTGTAGATCTTATTTGCTATTTTCCAAATTTATGACCCAGAAATAT----- 912
QY 334 GlyGlyIleGluValLysGlyTyrLysAsnGlnLeuThrArgGluIleTyrThrGln 353
Db 913 -----CCAAATGAGATTAATCTGAATCTTACCAAGAAAGTTAT--ACGAAT 957
QY 354 IleAsnPheAspArgLeuProGlnLeuArgValGlnProAsnLeuAlaThrMetGluTyr 373
Db 958 GTTATATCAGATACATTT-----AGACATATACAGAACTAGAAAT 999
QY 374 AsnLeuThrArgAlaSerPheLysLeuPheSerPheLeuGlnGlnPheIlePheTyrThr 393
Db 1000 GGATTTAACTAGA--AATCTCAATTTATTTACTTGATTAACCAAGGCGTTTTCACACA 1056
QY 394 GluAsnThrAsn-----PheGlyAsnArgLeu 402
Db 1057 AGAAATTTCTCGAGACATTTCTGATCTTATGATATTTTCTTTTACAGTACACGAGTG 1116
QY 403 ValGlyIleSerAsnArgAspAlaProThrTyrSerAsnThrIleThrGluThrLeuTyr 422
Db 1117 -----GCTTTACACATCTATATGATGATGATGATGATGATGATGATGATGATGAT 1155
QY 423 GlyIleArgThrGly-----SerProThrIleLysThrIleArgProPheGluSer 439
Db 1156 GGAGCGGTTATGATGATATTTATTTCTCAAGACATCCAAAGATTTCTTTTATATAGA 1215
QY 440 Tyr-----LysValSerIleValThrAspArgGlnSerProProValSerPro 455
Db 1216 AACAAACCTATGATTAAGGTGCAAAATGTGCAGACATAGAGATCTCAGATTAATA-- 1272
QY 456 IleGlnProHisPheIleIleAsnGlnIleGluLeuTyrLeuAsnGlySerSerAsn 475
Db 1273 -----TATGAAAGATATTTTTCGAATGACACTGA 1305
QY 476 ThrLeuLysTyrSerAlaGlyLysLeu-----SerAsnTyrGlnAsnThrThrPhePhe 494
Db 1306 GTATTTTCGATATTCATCCAAATTCACAAATATGAAATATTAATAAAGAACTGATCTTAT 1365
QY 495 GlnPheProArgLysLeuLysAspCysAsnLeuValIleAspProGlyCysSerProAsn 514
Db 1366 ATGATTTCCAAACAAACATGAAATAT----- 1392
QY 515 AsnAsnTyrSerHisIleLeuSerHisPheSerLeuPheThrTyrSerTyrValIleGly 534
Db 1393 AAAAGATATGCTCATATCTATCTATGATTAATAAAGCTAATATTAATTTTCAAGATTT 1452
QY 535 LeuGlnLeuGlnIleLeuAspThrGlyValLeuGlyTyrThrHisSerSerValAspArg 554
Db 1453 AGAGAAAGAAAGAGTT-----GCATTTAGTTGACACATCTAGTGTGATTTTC 1503
QY 555 TyrAsnAlaIleSerAspLysIleIleThrMetIleProAlaIleLysGlyAsnAsnLeu 574
Db 1504 CAATATACAAATGATTTATGATTAACATCACCCAAATCCAGCTCTAAACCTTTGAAGTA 1563
QY 575 AspThrAsnSerLysValIleGlyGlyProGlyHisIleThrGlyValAsnLeuValTyrLeu 594
Db 1564 AGTTTCTGATTCAAAAATTTGAAAGGCTCTGATCACAAGGAGGACCTTGATATCTTT 1623
QY 595 GlnSerGlnGlyArgLeuGluIleThrCysGluThrProAsnSerThrGlnSerTyrPhe 614
Db 1624 AAAAGATATGATGATTTATGATTTAGATTT--TTAAAAAATGTTTCTGCAAAATATCA 1680

```

QY 615 IleatgleuArgTyrAlaThranglYalaglYasnThrLeuProAsnIleSerLeuThr 634  
 Db 1681 GTACGTATTCGTTAGTACTACTAATGCT-----CCAAAGACAAACGATATTC 1725  
 QY 635 IleProGlyValIleGlyIleProProGlnArgLeuAsnThrPheSerGlyThrAsn 654  
 Db 1726 TTAACCGGAATAGATTAAGTGTGAG---CTCCCTTAGACCACTTCCCGCAAAAC 1782  
 QY 655 TyrAsn-----AsnLeuGlnTyrGlyAspPheGlyTyrPheGlnPheProSerThrVal 672  
 Db 1783 CCAATGCTACAGATTTAAACATATGCAGATTTTGGATATGTAACTTTCCAAAGAACGTT 1842  
 QY 673 -----ThrLeuProLeuAsnArgAsnIleProPheIle 683  
 Db 1843 CCAATTAACATTTGAAGAGAGAGACCTTATTAATGACC-----TTA 1887  
 QY 684 PheAsnArgAlaAspValSerAsnSerIleLeuIleIleAspValIleGluPheIlePro 703  
 Db 1888 TATGGTACACCAATCATTCATATATATA---TATATTGACAAATCGAATTTATTTCCA 1944  
 QY 704 IleThrSerSerMetHisGlnAsnArgGluGlyGlnLeuGluThrIleGlnThrIlys 723  
 Db 1945 ATCACTCAATCTGTATTTAGATTATACAGAGAGCAAAATATTAAGAAAAACACGAAAAATA 2004  
 QY 724 IleAsnThrPhePheThrAsnHisThrIysThrLeu 735  
 Db 2005 GTGAATGATTTATTTGTTAATTAACAAAGTTCTT 2040  
 RESULT 12  
 AA081180  
 ID AA081180 standard; DNA; 2061 BP.  
 AC AA081180;  
 DT 25-MAR-2003 (revised)  
 DT 12-AUG-1995 (first entry)  
 XX B.C. toxin PST1M3 gene.  
 DE  
 XX Delta-endotoxin; crystal protein; biological control agent; Calliphoridae;  
 KW screw-worm; sheep blowfly; Lucilia; Phormia; Calliphora; insecticide;  
 KM pesticide; B.C.; 86.  
 XX  
 Os Bacillus thuringiensis.  
 XX  
 PN MO502694-A2.  
 XX  
 PD 26-JAN-1995.  
 XX  
 PF 13-JUL-1994; 94WO-US007902.  
 XX  
 PR 15-JUL-1993; 93US-00093199.  
 PA (MYCO) MYCOGEN CORP.  
 XX  
 PI Hickie LA, Payne J;  
 XX MPI; 1995-067338/09.  
 DR P-P8DB; AAR63079.  
 XX  
 PT Method for controlling Calliphoridae pests - specifically utilises  
 FT Bacillus thuringiensis isolates or toxins.  
 XX  
 PS Disclosure; Page 42-43; 50pp; English.  
 CC A library was constructed from Bacillus thuringiensis PST1M3 total  
 CC cellular DNA in lambda Gem-11. Plasmid pMYC1636, selected in Escherichia  
 CC coli, contained a 15 kb insert expressing a beta-endotoxin gene. This was  
 CC sequenced (AA081180). A cured, acrysaliferous B.C. host carrying  
 CC pMYC1636 produced a 77 kDa crystal protein (AAR63079). (Updated on 25-MAR  
 CC -2003 to correct PN field.)  
 XX

SQ Sequence 2061 BP; 779 A; 331 C; 319 G; 632 T; 0 U; 0 Other;  
 Alignment Scores:  
 Pred. No.: 3,866-71 Length: 2061  
 Score: 899.50 Matches: 247  
 Percent Similarity: 48.38 Conservative: 126  
 Best Local Similarity: 32.08 Mismatches: 270  
 Query Match: 23.24 Indels: 129  
 DB: 2 Gaps: 30  
 US-10-783-417-2 (1-735) x AA081180 (1-2061)  
 QY 1 MetAsnGlnAsnAspAsnAsnGlnTyrGluIleIleAspSerHisThrSerProTyr 20  
 Db 1 ATGAATCATATTCAAATATGAAATGAAATGAAATATTAATCAATGCTCCATCCATGATTTT 60  
 QY 21 PheProAsnArgAsnSerAsnAspSerGlyTyrProTyrThrAsnAspProAsnGlnPro 40  
 Db 61 -----ACCAAGTCTAATTAATCTATTTAGATATTCATTAAGCAAAATGACCAATCAACCA 114  
 QY 41 LeuGlnAsnThrAsnTyrGlyGluTyrPheAsnMetCysGlnGlyAsnThrGlnTyrGly 60  
 Db 115 CTGAAAAACAGATTAACAAAGATTTGCTCATGTGTCTCAAGATATCAACAAATATGAC 174  
 QY 61 AspAsnPheGluThrPheAlaSerAlaAspThrIleAlaAlaValSerAlaGlyThrIle 80  
 Db 175 AATAATCGCGGGAATTTTGTAGTTCTGAACATATGTTGAGATGAGTACAGATATTTAT 234  
 QY 81 ValSerGlyThrLeuLeuAlaGlyIleGlyIleLeuThrSerIleSerGlyProIleGly 100  
 Db 235 GTAGTAGGAATCACTGTTAGCA-----GCTTTGCTGCCCT----- 270  
 QY 101 IleIleGlyAlaIleIleIleSerPheGlyThrLeuIleThrValPheTyrProAlaGly 120  
 Db 271 GTCTTAGCTGACGATATATATCTTTGGACCTTTGTCGCCGATTTTGG---CAAGCA 327  
 QY 121 GluGlnAspIysThrValTyrThrGlnPheIleIleIleIleIleIleIlePheValAspThr 140  
 Db 328 TCTGACCTCGCAAAATGTTGGCAGGATTTGTAACATCGCA-----GGAAG 375  
 QY 141 ProLeuThrGlu-----SerIleGlyGlnLeuIleGlnThrLeuGlnIleGlyPhe 157  
 Db 376 CCTATACAGAAATAGATTAACAAACATTAATTAATGACTTAACCTTCTATGTAACACTATA 435  
 QY 158 ArgGlnIleLeuGlnSerTyrAsnThrAlaLeuAspAspTyrPargIysLeuIleAspGlu 177  
 Db 436 AAAAATCAACTTGAATATATCAAGAATTTTGCATTAATGAGCCAGCACGT----- 489  
 QY 178 GlnAlaProGlyLeuProProSerSerAlaLeuGlnGlnAlaAlaLeuThrLeuIleIle 197  
 Db 490 -----ACACACGCTAATGCTTAAGCAGACATGATGATCTTTACT 528  
 QY 198 ArgPheGluAsnValHisAsn-----AspPheIleArgGluIleProGlyPheGln 214  
 Db 529 ACCTTAGAAGCTTAATATAGATTAAGAAATTAATTAATTAATTAATGCTAGATATGCA 588  
 QY 215 LeuGluThrTyrIysThrLeuLeuLeuProIleTyrAlaGlnAlaAlaAsnPheHisLeu 234  
 Db 589 ATACCAACA-----CTCCCTGATATGCAACAATAGCTACTTGGACATTTG 633  
 QY 235 AsnLeuLeuGlnGlnGlyAlaGluLeuAlaAspGluTyrPheAlaAspIleHisProSer 254  
 Db 634 AATTTATTAACATGCTGCTACTATTAACATATATG-----CTGCAAAATCA 684  
 QY 255 GlnIleGluProAsnAlaGlyThrSerAspAspTyrTyrIys---LeuLeuIleGluAsn 273  
 Db 685 GGTATTAATCCAAAGTACTTTCAATTCATCTAATTAATCAAGGCTATTTAAACGTAA 744  
 QY 274 IleProIysTyrSerAsnTyrCysAlaAsnThrTyrArgThrGlyLeuIleAsnLeuArg 293  
 Db 745 ATCAAGAAATATCTGACTATTTGTATACAAACGATCAATGACGACTACTATATTTGA 804  
 QY 294 AspGluProAsnMetIysTyrSerIlePheAsnAspTyrArgArgTyrMetThrIleThr 313

Db	805	ACTATATCTAACGCAACATGGAATATGTAATACTACTACCGTTTAGAAATGACTTAACT	864
Oy	314	ValleuAapThrlIeSerGlnPheSerLeuTyraPilleTyarsGlyrArGaspSerIle	333
Db	865	GTGTATGATCTTAATGCTATATTTTCCAAATATATGACCCAGAAAATAT-----	912
Oy	334	GlyGylIleGluValValGylIleIleYasnGlnLeuThraParglIleTyThrIu	353
Db	913	-----CCAAATGAGATTAAATCTGAACCTTACCAAGAAAGTTTAT---ACGAAT	957
Oy	354	IleAsnPheAparGleuProGlnLeuArgValGlnProAsnLeuAlaThrMetGluTy	373
Db	958	GTTAATTCAGATACCTT-----AGAACCTTAACCAAGACTGAANAAT	999
Oy	374	AsnLeuThraArgAlaSerPheIleuPheSerPheLeuGlnIlePheIleTyThr	393
Db	1000	GGATTAACTAGA---AATCTCAATATTATTCTTGGAATAAACCAAGGCGTTTTTACCA	1056
Oy	394	GluAenThraAsn-----PheGlyAsnArgLeu	402
Db	1057	AGAAATTCAGACATCTTGATCCTTATGATATTTTTCTTTTACAGGTACAGAT	1116
Oy	403	ValGylIleSerAsnArgAlaProThrlTySerAsnThrlIeThGluThrLeuTy	422
Db	1117	-----GCCCTTACACATCTAAATGATATGATCGCAATATATCGG	1155
Oy	423	GlyGluArgThrgly-----SerProThrlTySerThrlIeArgProPheGluSer	439
Db	1156	GGAGCGGCTTACGACATATATTTCTCAAGACACATCCAAAGATATTCCTTTTATAGA	1215
Oy	440	Tyr-----IleValSerIleValThraPargIlnSerProProValSerPro	455
Db	1216	AACAAACCTATGATTAAGTTCGAATTTCTCAGACATAGAGTCACTCAGATATATA--	1272
Oy	456	IleGlnProIlePheIleIleAsnGlnIleGluLeuTyrlleuAsnGlySerSerAsn	475
Db	1273	-----TATATAATGATATTTTTTCCAAATAGCAGTAA	1305
Oy	476	ThrlleuTyrlSerAlaGlyIleSerLeu---SerAsnTyrlGlnAsnThrlIePhe	494
Db	1306	GTAATTCGATTTATCATTCCAATTCACATATAGAAATATTTTAAAGAACTGATCTTAT	1365
Oy	495	GlnPheProArgIleYAspCysAsnLeuValIleAspProGlyCysSerProAsn	514
Db	1366	ATGATTTCCAAACCAACATGAGAAAAT-----	1392
Oy	515	AsnAsnTyrlSerHisIleLeuSerHisPheSerLeuPheThrlTySerTyValIleGly	534
Db	1393	AAAGAAATATGGCTACTCTATCGATATTAATAAACTGATTAATTAATTTTCAGTGT	1452
Oy	535	LeuGlnLeuGlnIleLeuAspThrglyValLeuGlyTrpThrlIeSerSerValAsp	554
Db	1453	AGAGAAAGAGAAAGATT-----GCATTTAGTTGAGACACATCACTAGTGTGATTC	1503
Oy	555	TyrAsnAlaIleSerAspIleIleIleThrlleProAlaIleTyGlyAsnAsnLeu	574
Db	1504	CAAAATACAAATATGATTAATGATTAATCACCAAATTCACGCTTAAAAAGCTTGAAGTA	1563
Oy	575	AspThraSerIleValIleGlnGlyProGlyHisThrglyGlyAsnLeuValTyrlleu	594
Db	1564	AGTTCGTATTCGAAATATGTGAAAGGCTCTGTCACACAGGTGAGACTTGATATCTT	1623
Oy	595	GlnSerGlnGlyArgLeuGlnIleThrcCysGlnThrProAsnSerThrlIeSerTyPhe	614
Db	1624	AAAGATAGTATGATTTTAGATTAAGATT---TTAAAAATATGTTCTCCGCAATATCAA	1680
Oy	615	IleArgLeuArgTyrlaIleThraSngIyAlaGlyAsnThrlleuProAsnIleSerLeuTh	634
Db	1681	GTAGCTATTCGTATATGCTAAATGCT-----CCAAAGACAAACGATATTC	1725
Oy	635	IleProGlyValIleGlyIleProProGlnArgLeuAsnThrlIeSerGlyThraAsn	654

Df	1726	TTAACCGGAAATAGTACTATAAAGTGTCGAG---CTCCCTAGTACCACTTCCCGCAAAC	1782
Qy	655	TyrAsn-----AsnLeuGlnIYrGIyaSPheGlyTYrPheGlnPheProSerThVal	672
Df	1783	CCAATGTCTACAGATTTAACATATATGCAGATTTTGGATATGTAAcATTGCCAAGAACAgtT	1842
Qy	673	-----ThleuProLeuasnArganlleProPheIle	683
Df	1843	CCAAATTAACCATTTGAAGAGAGACACCTTTATTAATGACC-----TTA	1887
Qy	684	PheasnaRglaaSPValSerAsmSerilleuIleleasPlysillegluPheilePro	703
Df	1888	TATGCTACCAcCAAAATCATTTCATATTAATATN---TATATTAACAANAATCGAATTTATTCCA	1944
Qy	704	IleThSerSerWeethisglinaeNaRgylLysGlnLysleugluthrlileGlnThrLys	723
Df	1945	ATCACTCAATCTGTGAATTAAGATTATACAGAGAAACAAAATATATGAAAAAACACAGAAATA	2004
Qy	724	IleasnThrPhePheThraamhlsthrLyThleu	735
Df	2005	GtGAATGATTATTTATTTATTAATAAACAAAGTTCTT	2040

RESULT 13  
 AAN50525 standard; DNA, 3756 BP.

XX	AAN50525;
DT	24-OCT-2003 (revised)
DT	25-MAR-2003 (revised)
DT	23-OCT-1991 (first entry)
DX	Bacillus thuringiensis var. israelensis endotoxin insert in plasmid pSY367.
DE	Endotoxin; insecticide; ss.
XX	
OS	Bacillus thuringiensis serovar israelensis.
XX	
PN	EP153166-A.
PD	28-AUG-1985.
PF	15-FEB-1985; 85EP-00301017.
PR	22-FEB-1984; 84US-00582506.
PR	22-JAN-1985; 85US-00693556.
PA	(SYTR ) SYNTRO CORP.
PI	Walfield AM, Pollock TJ;
DR	WPI; 1985-211724/35.
PT	Polypeptide active against Diptera insects - prep'd. from DNA sequence coding for BTI endotoxin using bacterial host.
PS	Disclosure; Page 15a-e; 27pp; English.
CC	The B. thuringiensis var. israelensis endotoxin insert in pSY367 is expressed in a bacterial host. The protein produced has insecticidal activity against dipteran insects. (Updated on 25-MAR-2003 to correct PA field.) (Updated on 24-OCT-2003 to standardise OS field)
SQ	Sequence 3756 BP; 1379 A; 566 C; 659 G; 1150 T; 0 U; 2 Other;

Alignment Scores:

Pred. No.:	4..2e-69	Length:	3756
Score:	881.00	Matches:	248
Percent Similarity:	47.4%	Conservative:	120
Best Local Similarity:	32.0%	Mismatches:	272
Query Match:	22.8%	Indels:	136
DB:	1	Gaps:	31

```

US-10-783-417-2 (1-735) x AAN50525 (1-3756)
QY 1 MetcAnglnAsnAspAsnAsnAngluYrGluIlelleAspSerHisThrSerProTyr 20
DB 941 ATGAATCCATATCAAAATAGAAATGAAATATGAAATATTCATATGCTCCATCCAAATGGTTT 1000
QY 21 PheProAsnArgAsnSerAsnAspSerArgTyrProTyrThrAsnAsnProAsnGlnPro 40
DB 1001 -----AGCAAGCTAAATATCTATCTTACATATCCATAGCAAAATAGCAAAATCAACA 1054
QY 41 LeuGlnAsnThrAsnTyrIleGluIlePleuAsnMetCysGlnGlyAsnThrGlnTyrGly 60
DB 1055 CTGAAACAAACACAAATTTACAAAGATTTGCTCAATGCTGTCAAGATATATCAAAATATGCG 1114
QY 61 AspAsnPheGluThrPheAlaSerAlaAspThrIleAlaAlaValSerAlaGlyThrIle 80
DB 1115 AATTAATGCGGGGAATTTCTGATCTGAACTATTTGAGTTAGTCAGATATATTT 1174
QY 81 ValSerGlyThrLeuLeuAlaGlyIleGlyGlyLeuThrSerIleSerGlyProIleGly 100
DB 1175 GTAGTAGGAATCTATGTAGGA-----GCTTTGCTGCCCT----- 1210
QY 101 IleIleGlyAlaIleIleIleSerPheGlyThrLeuIleThrValPheTrpProAlaGly 120
DB 1211 GCTTAGCTGACAGATATATATCTTTGGACTTTGTTGCCGATCTTTGG---CAAGGA 1267
QY 121 GluGlnAspLysThrValTrpThrGlnPheIleLysMetGlyGluIlePheValAspThr 140
DB 1268 TCTGACCTCGCAAAATGTTGGCAGATTTGTTAAACATCGA-----GGAAGG 1315
QY 141 ProLeuThrGlu-----SerIleLysGlnLeuLysLeuGlnThrLeuGlnGlyPhe 157
DB 1316 CCTATCAAGAAATATAGATATAAAACATTAATATGACTTAATCTTATGTAACACTTANA 1375
QY 158 ArgGlnIleLeuGlnSerTyrAsnThrAlaLeuAspAspTrpArgLysLeuLysArgLeu 177
DB 1376 AAAAATCAACTGATTAATATACAGAAATTTTCGATAAATGGAGCCAGCACAGT----- 1429
QY 178 GlnAlaProGlyLeuProProSerSerAlaLeuGlnGlnAlaAlaLeuThrLeuLysIle 197
DB 1430 -----ACACACGCTAATGCTAAAGCAGTACATATCTTTTACT 1468
QY 198 ArgPheGluAsnValHisAsn-----AspPheIleArgGluIleProGlyPheGln 214
DB 1469 ACCTTAGAAGCTATATATAGATAAGATTAGATATGTTAAAAAATATATGCTATGCA 1528
QY 215 LeuGluThrTyrLysThrLeuLeuLeuProIleTyrAlaGlnAlaAlaAsnPheHisLeu 234
DB 1529 ATACCAACA-----CTCCCTGCATATGACAAATATGCTACTTGGCAGCTTG 1573
QY 235 AsnLeuLeuGlnGlnGlyValaGluLeuAlaAspGluTrpAsnAlaAspIleHisProSer 254
DB 1574 AATTATATTAACATGCTGCTACTCTTACATTAACAAATATGGA-----CTGCAAAATCA 1624
QY 255 GlnIleGluProAsnAlaGlyThrSerAspAspTyrTyrLys-----LeuLeuLysGluAsn 273
DB 1625 GGATTAATATCAAGATCTTCAATTCATCTAATTAATCTATCAGGGCATTTAAAAAGCTAA 1684
QY 274 IleProLysTyrSerAsnTyrCysAlaAsnThrTyrArgThrGlyLeuLysAsnLeuArg 293
DB 1685 ATACAGAAATATATCTGATATGTATATACAAACGTACAAATGACAGACTAATGATTTTAA 1744
QY 294 AspGluProAsnMetLysTrpSerIlePheAsnAspTyrArgArgTyrMetThrIleThr 313
DB 1745 ACTAATATCTAACGCAACATGAAATATGATTAATCTTACCGTTTGAATAATGACTTAAT 1804
QY 314 ValLeuAspThrIleSerGlnPheSerLeuTyrAspIleLysArgTyrArgAspSerIle 333
DB 1805 GGTATATGATCTTATGCTATTTTCCAAATATATGACCCAGAAAATAT----- 1852
QY 334 GlyGlyIleGluValLysGlyIleLysAsnGluLeuThrArgGluIleTyrThrThrGlu 353
DB 1853 -----CCAAATGAGATTAATCTGAACTTATACAGAAAGTTAT---ACGAAT 1897
QY 354 IleAsnPheAspArgLeuProGlnLeuArgValGlnProAsnLeuAlaThrMetGluTyr 373
DB 1898 GTTAAITGATATACATTT-----AGAACATTAACAGAACTAGAAAAT 1939
QY 374 AsnLeuThrArgAlaSerPheLysLeuPheSerPheLeuGlnGlnPheIlePheTyrThr 393
DB 1940 GGATTAATACATGA---AATCTTACATTAATTAATCTTGATTAACCAAGGCGCTTTTACACA 1996
QY 394 GluAsnThrAsn-----PheGlyAsnArgLeu 402
DB 1997 AGAAATTTCTCGACATCTTGATCTTATGATATTTTCTTTTACAGGTAAACAGATG 2056
QY 403 ValGlyIleSerAsnArgAspAlaProThrTyrSerAsnThrIleThrGluThrLeuTyr 422
DB 2057 -----GCTTTACACATTAATATGATATGATCGCAACATATCTCG 2095
QY 423 GlyIleArgThrGly-----SerProThrThrLysThrIleArgProPheGluSer 439
DB 2096 GGAGCGGTTCAATGAAATATTAATTTCTCAAGACATCCAAAGATATTTCTTTTATAGA 2155
QY 440 Tyr-----LysValSerIleValThrAspArgGlnSerProProValSerPro 455
DB 2156 AACAACTTATGATTAAGCTGAAATTTGTGACACATAGAGACTACTCAGATTAATA--- 2212
QY 456 IleGlnProHisPheIleIleAsnGlnIleGluLeuTyrLeuAsnGlySerSerAsnAsn 475
DB 2213 -----TATGAAATGATATTTTTCGAATATGACGTGAA 2245
QY 476 ThrLeuLysTyrSerAlaGlyLysLeu-----SerAsnTyrGlnAsnThrThrPhePhe 494
DB 2246 GTATTCGATTAATTCATCAATTCACAAATAGCAAAATTAATTAATAAGAACTGATCTTAT 2305
QY 495 GlnPheProArgLysLysAspCysAsnLeuValIleAspProGlyCysSerProAsnPhe 514
DB 2306 ATGATTTCAAAACAAACATGAAATAT----- 2332
QY 515 AsnAsnTyrSerHisIleLeuSerHisPheSerLeuPheThrTyrSerTyrValIleGly 534
DB 2333 CAAGAAATATGCTACATCTATCTATGATATAAAACATGATTAATTAATTTCAAGTAT 2392
QY 535 LeuGlnLeuGlnIleLeuAspThrGlyValLeuGlyTyrProHisSerSerValAspArg 554
DB 2393 AGAGAAAGAAAGAAAGAGTT-----GCATTAAGTTGACACATATGATGTTATTTTC 2443
QY 555 TyrAsnAlaIleSerAspLysIleIleThrMetIleProAlaIleLysGlyLysAsnLeu 574
DB 2444 CAAAATTAACATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2503
QY 575 AspThrAsnSerLysValIleGluGlyProGlyHisThrGlyGlyAsnLeuValTyrLeu 594
DB 2504 AGTTCTAATTTCAAAATTTGTGAAGCCCTCGTCAACAAAGAAACCTTGATATCTT 2563
QY 595 GlnSerGlnGlyArgLeuGluIleThrCysGluThrProAsnSerThrGlnSerTyrPhe 614
DB 2564 AAAAGATAGATGAATTTTAGAGTTAGATT---TAAAAAAAGTTTCCAAACATATCA 2620
QY 615 IleArgLeuArgTyrAlaIleThrAsnGlyAlaGlyAsnThr-----LeuPro 629
DB 2621 GTACGTATTTGTTATGCTAATAGCTCCAAAGACAGATTTCTTAAACCGAAATAGAT 2680
QY 630 AsnIleSerLeuThrIleProGlyValIleGlyIleProProGlnArgLeuAsnAsnThr 649
DB 2681 ACTATTAAGTTGAGCTCCCT-----AGTACC 2707
QY 650 PheSerGlyThrAsnTyrAsn-----AsnLeuGlnTyrGlyAspPheGlyTyrPheGln 667
DB 2708 ACTTCCCGCAAAACCAAAATGCTACAGATTTTAACATATGACGATTTTGATATAGTAA 2767
QY 668 PheProSerThrValThrLeuProLeuAsnArgAsnIle-----ProPheIle 683
DB 2768 TTTTCAGAAACAGTTTCA-----AATTAACATTTGAAGAGAAACACTTTTAT 2818

```

QY 684 -----PheAnArg1AAspValSerAemSerIleuIleIleAspIysIle 699  
DB 2819 AATGACCTTATTATGTAACCAAAATCTTCAATATATATA---TATATTGACAAAAAT 2875  
QY 700 GIupheIleProIleTherSerMetH1sgInAsnArgIuLysGlnLysLeuThr 719  
DB 2876 GAATTTATTCATCACTCAATCTGTATTAAGATTATACAGAGACAAATATAGAAA 2935  
QY 720 IlegInThrLysIleAsnThrPhePheThrAsnH1stHlyThrLeu 735  
DB 2936 ACAAGAAATAGTGAATGATTATTATTAAATTAACAAAGTTCTT 2983  
RESULT 14  
AAD43974  
ID AAD43974 standard; DNA; 4896 BP.  
XX AAD43974;  
AC XX  
XX  
DT 13-DEC-2002 (first entry)  
DS Bacillus thuringiensis ssp. finitimus cry28Aa1 gene.  
KM Delta-endotoxin; cry26Aa1; insect-resistant plant; toxin;  
KW transgenic host cell; insecticide; gene; ds.  
XX  
OS Bacillus thuringiensis.  
XX  
XX Key Location/Qualifiers  
FH CDS 1129..4458  
FT /\*tag= a  
FT /product= "cry28Aa1 gene"  
XX  
XX US2002038005-A1.  
XX  
XX 28-MAR-2002.  
XX  
XX 08-JAN-2001; 2001US-00756526.  
XX  
XX 07-JAN-2000; 2000US-0175158P.  
XX  
XX (WOJC/) WOJCIECHOWSKA J A.  
XX (LEWI/) LEWITTIN E I.  
XX (ZALU/) ZALUMIN I A.  
XX (REVI/) REVINA L P.  
XX (CHES/) CHESTUKHINA G G.  
XX  
XX Wojciechowska JA, Lewitin EI, Zalumin IA, Revina LP;  
XX Chestukhina GG;  
XX  
XX WPI; 2002-403936/43.  
XX P-PSDB; AAE26353.  
XX  
XX Novel isolated delta-endotoxin nucleic acid molecules, cry26Aa1 and  
XX cry28Aa1 isolated from Bacillus thuringiensis finitimus, that encodes  
XX toxin active against insects, useful for controlling insects.  
XX  
XX Claim 1; Page 33-37; 42pp; English.  
XX  
XX The invention relates to isolated delta-endotoxin nucleic acid molecules,  
XX cry26Aa1 and cry28Aa1 isolated from Bacillus thuringiensis finitimus,  
XX that encode a toxin that is active against insects. The invention is  
XX useful for producing an insect-resistant plant, by introducing the  
XX nucleic acid molecule into the plant, where the nucleic acid is  
XX expressible in the plant in an effective amount to control an insect. The  
XX invention is useful for producing a toxin that is active against insects  
XX by obtaining the transgenic host cell and expressing the nucleic acid  
XX molecule in the host cell, which results in the toxin that is active  
XX against insects. The toxin is useful for controlling an insect by  
XX delivering to the insect an effective amount of toxin. The invention is  
XX useful for controlling insects. The toxin is useful for inhibiting the  
XX ability of insect pest to survive, grow or reproduce, for limiting insect  
XX -related damage or loss in crop plants, and to prophylactically treat

CC insect susceptible areas or plants to confer protection or resistance  
CC against harmful insects. The present sequence is Bacillus thuringiensis  
CC ssp. finitimus cry28Aa1 gene  
XX  
SQ Sequence 4896 BP; 1727 A; 757 C; 887 G; 1525 T; 0 U; 0 Other;  
Alignment Scores:  
Pred. No.: 4.8e-66 Length: 4896  
Score: 849.00 Matches: 229  
Percent Similarity: 48.8% Conservative: 106  
Best Local Similarity: 33.3% Mismatches: 248  
Query Match: 21.9% Indels: 104  
DB: Gaps: 25  
US-10-783-417-2 (1-735) x AAD43974 (1-4896)  
QY 68 SerAlaAspThrIleAlaIleValSerAlaGlyThrIleValSerGlyThrLeuAla 87  
DB 1210 AGCAGTGATACAGCTGCTGTAGTAAGCCAGGAGATTGTGGGTACTATACGACA 1269  
QY 88 GlyIleGlyLysLeuTherSerIleSerGlyProIleGlyIleIleGlyAlaIleIle 107  
DB 1270 -----GCCTTTCATCATTTGTTATCA-----GGTGTGACTATATA 1308  
QY 108 SerPheGlyThrLeuIleThrValPheTrpProAlaGlyGlnAsp---LysThrVal 126  
DB 1309 TCATTTGACACTTGTGCTCCGCTTCTTGCGCTGATCCAGAGAAATCCAAAAAATT 1368  
QY 127 TrpThrGlnPheIleLysMetGlyGluIlePheValAspThrProLeuThrGluSerIle 146  
DB 1369 TGGTCACATTTATGAAACACGAGAGACCTTTAAATCAAAATTTCTACAGCTGTA 1428  
QY 147 LysGlnLeuLysLeuGlnThrLeuGluGlyPheArgGlnIleLeuGlnSerIleThr 166  
DB 1429 AAAAGAAATGACATTCATCTCAATTAATGTTTAAAGATTTAAGCTACTATGAAGA 1488  
QY 167 AlaLeuAspAspTrpArgLysLeuLysArgLeuGlnAlaProGlyLeuProProSerSer 186  
DB 1489 GCATTTAATGATTTGAAGAAGA-----AATCCAACTGCA 1521  
QY 187 AlaLeuGlnGlnAlaIleLeuThrLeuLysIleArgPheGluAsnValHisAsnAspPhe 206  
DB 1522 -----AATACGCCAGATGGTATGATCAGAGATTTGAAGCGCTCATTTCAATTT 1572  
QY 207 IleArgGluIleProGlyPheGlnLeuGluThrIleLysThrLeuLeuLeuProIleIle 226  
DB 1573 GTACCAATATGCCCAACTCCCACTCCACGTATGACACATTTATTATTAAGTGTCTAT 1632  
QY 227 AlaGlnAlaIleAsnPheHisLeuAsnLeuLeuGlnGlnIleValGluLeuAlaAspGlu 246  
DB 1633 ACAAGAGCTGCAAAATTTATCAATTTGATTAATCAATCAAGGTGACAAATTCGCGGATCA 1692  
QY 247 TrpAsnAlaAspIleHisProSerGlnIleGluProAsnAlaGlyThrSerAspAspIle 266  
DB 1693 TCGATTCAGATCAACACCAATTCACCAATGTTGAAGTATCATCGATCACT-----TAT 1743  
QY 267 TyrLysLeuLeuLysGluAsnIleProLysIleYserAsnIleCysAlaAsnThrIleArg 286  
DB 1744 TATGACGAGCTATTTGTTATTTATTAAGTATTAATTTATTTGACCAAGACATTCAT 1803  
QY 287 ThrGlyLeuLysAsnLeuArgAspGluProAsnMetLysTrpSerIlePheAsnAspIle 306  
DB 1804 AAGAGTTGAATCACCTTAAAGAAATCAGAAAAATCAACATGGAGATCTTATACACAT 1863  
QY 307 ArgArgIleMetThrIleThrValLeuAspThrIleSerGlnPheSerLeuIleIle 326  
DB 1864 CGTCGAGAAATGACCTTATGTATGTATGATCTTTCGCAACTTTCCTTTATATATATA 1923  
QY 327 LysArgIleArgAspSerIleGlyGlyIleGluValLysGlyIleLysAsnGluLeuThr 346  
DB 1924 CCTCGTTTT-----CCAGAGGAGTAGAATCAATTAATTAACA 1959  
QY 347 ArgGluIleIleThrThrGluIleAsnPheAspArgLeuProGlnLeuArgValGlnPro 366

```

Db      1960 AGAGAGGTTTATACAAAGTTTAGATCATTTAAACACACACCA----- 2001
Qy      367 AenleuAlaThrMetGluTyrAsnleuThrArgAlaSerPheLeuPheSerPheLeu 386
Db      2002 -----GGGCTATTATCTGGCTG 2019
Qy      387 GluGlnPheIlePheTyrThrGluAsnThrAsnPhelGlyAsnArgLeuValGlyIleSer 406
Db      2020 TCAGATATTGAGTTTACACGAGAGGTGGCAGAAAGCATTTATTTATTCAGGTATTT--- 2076
Qy      407 AsnArgAlaProThrTyrSerAsnThr-----IleThrGluThrLeuTyrGly 423
Db      2077 ---CGAGAGTCTAAATATTATTAAGTGTATCAATTTTATACGATGAAATATTTATGCT 2133
Qy      424 GluArgThrGly---SerProThrThrIleArgProPheGluSerTyrIleVal 442
Db      2134 AATACAAATAGATTAGTAAAGAGCTCATTTATTTACCGAGGAA---TTTATGACT 2190
Qy      443 SerIleValThrAspArgGlnSerProValSerProIleGlnProIlePheIleIle 462
Db      2191 CACTTAAGCATTAACCGTCTTTCAAAACAATAGCTGGATTAATATAGTTTACAGTTTA 2250
Qy      463 AsnGlnIleGluLeuTyrLeuAsnGlySerSerAsnAsnThrLeuLeuTyrSerAlaGly 482
Db      2251 ATTCAAAAATCGTATTCACAACTTTTAAACGATAT---GAATATCAAAAAAAT 2304
Qy      483 GlySerLeuSerAsnTyr-----GlnAsnThrThrPheGlnPheProArgLys 499
Db      2305 TTTATGTGATTAATCAAAAATGAACTTCAAGAACTTAC----- 2343
Qy      500 LysAspCysAsnLeuValIleAspProGlyCysSerProAsnPheAsnAsnTyrSerHis 519
Db      2344 ---AACTATCTTAATGATTTATGTGGTCA---AACAGCCAAAATTAACAAT 2391
Qy      520 IleLeuSerHisPheSerLeuPheThrTyrSerTyrValIleGlyLeuGlnLeuGlnIle 539
Db      2392 AATTATCTCATTTTCCATTTATCATCCAC-----AAGTTACAGTTT 2433
Qy      540 LeuAsp-----ThrGlyValLeuGlyTyrThrHisSerSerValAspArgLys 555
Db      2434 GCTGAGTATTTTCACTCATATTTTGCATTAGGTTGACACACAAATAGTTAACTCCCA 2493
Qy      556 AsnAlaIleSerAspLysIleIleThrMetIleProAlaIleIleGlyAsnAsnLeuAsp 575
Db      2494 AATTAAATATCAGAAAGTGTGATACAAATCCATTGGTAAAGCTTACGAAGTT--- 2550
Qy      576 ThrAsnSerLysValIleGlyIleProGlyHisThrGlyGlyAsnLeuValTyrLeuGln 595
Db      2551 ACTAACAAATTCAGTTATTAAGAGACAGGTTTACAGGTGGAGATTATTAATGAACCTTCT 2610
Qy      596 SerGlnGlyArgLeuGluIleThrCysGluThrProAsnSerThrGlnSerTyrPheIle 615
Db      2611 GAT-----AAATGTTCTATTAAATGTAAGCT---AGTTCTTAAAAAAAAGCTTATA 2661
Qy      616 ArgLeuArgTyrAlaThrAsnGlyValGlyAsnThrLeuProAsnIleSerLeuThrIle 635
Db      2662 AGTCAATTTTATGCTGCAAAATAGCAATAGCTGATCAATAGACGTAGCTATTCGGA 2721
Qy      636 ProGlyValIleGlyIleProProGlnArgLeuAsnAsnThrPheSerGlyThrAsnTyr 655
Db      2722 GCAGAGAGTTCTA-----TTGCAACCTACCTTTTCTAGAAAGAGGAAC 2763
Qy      656 AsnAsn-----LeuGlnTyrGlyAspPheGlyTyrPheGlnPheProSer 670
Db      2764 AATAAATTTTACAAATTCAGACCTTAATAGATTAAGATTTTCAATATCATACCTTTAGCT 2823
Qy      671 ThrValThrLeuProLeuAsnArgAsnIleProPheIlePheAsnArgAlaAsp----- 688
Db      2824 GATATTGAATTAATCCGAAAGTGAAGAATTCATATCCATTGCAACCGAGAGATGATAT 2883
Qy      689 ValSerAsnSerIleLeuIleIleAspLysIleGluPheIleProIleThrSerSerMet 708

```

```

Db      2884 GAGAGGAGTGAATCTTTTAAATGATTAATTAGATTCAAACTTAGATGAAAAATTAT 2943
Qy      709 HisGlnAsnArgGlyLeuGlnLeuGluThrIleGlnThrLysIleAsnThrPhePhe 728
Db      2944 -----ACTAATGAATGAATTTAGAGAGCAAGCAAGAAAGCATGATATTTATTT 2994
Qy      729 ThrAsnHisThrLysThrLeu 735
Db      2995 ATAAACGCACAAACGCTTTG 3015

RESULT 15
ADP31301
ID      ADP31301 standard; DNA; 4896 BP.
AC      ADF31301;
XX      12-FEB-2004 (first entry)
DE      Bacillus thuringiensis serovar finitimus pf2 DNA clone.
XX      Cry26Aa1; cry28Aa1; delta-endotoxin; insect pest control;
XX      transgenic plant; insect resistance; insecticide; gene; ds.
OS      Bacillus thuringiensis serovar finitimus.
XX      Key      Location/Qualifiers
FH      CDS      1129..4458
FT      /tag= a
FT      /product= "Cry28Aa1 delta-endotoxin protein"

EN      US2003150018-A1.
XX      07-AUG-2003.
PD      15-JAN-2003; 2003US-00345020.
XX      07-JAN-2000; 2000US-0175158P.
PR      08-JAN-2001; 2001US-00756526.
XX      PA      (WOJC/) WOJCIECHOWSKA J A.
XX      PA      (LEMI/) LEWITIN E I.
XX      PA      (ZALU/) ZALUNIN I A.
XX      PA      (REVI/) REVINA L P.
XX      PA      (CHES/) CHESTUKHINA G G.
PI      Wojciechowska JA, Lewitin EI, Zalunin IA, Revina LP;
PI      Chestukhina GS;
XX      DR      WP1; 2003-897623/82.
XX      DR      P-PSDB; ADF31302.
XX      PT      New isolated nucleic acid molecule encoding a toxin that is active
XX      PT      against insects useful for controlling insect pests or for conferring
XX      PT      insect resistance in plants.
XX      PS      Claim 7; SEQ ID NO 3; 42pp; English.
XX      CC      The present invention relates to the isolation of novel cry26Aa1 and
XX      CC      cry28Aa1 delta-endotoxin genes from Bacillus thuringiensis sep.
XX      CC      finitimus. The sequences for the delta-endotoxin polypeptides are also
XX      CC      disclosed. The invention provides methods for producing the toxins and
XX      CC      compositions containing the toxins. The methods and sequences of the
XX      CC      invention are useful for controlling insect pests in transgenic plants to
XX      CC      confer insect resistance. The present sequence represents a DNA clone
XX      CC      that contains the coding sequence for cry28Aa1 delta-endotoxin.
SQ      Sequence 4896 BP; 1727 A; 757 C; 887 G; 1525 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:      4.8e-66      Length:      4896
Score:          849.00      Matches:      229
Percent Similarity: 48.8%      Conservative: 106
Best Local Similarity: 33.3%      Mismatches: 248

```



Query Match:	21.9%	Indel:	104
DB:	10	Gaps:	25
US-10-763-417-2 (1-735) x ADF31301 (1-4896)			
Qy	68	SerialAspThrIleAlaIValSerAlaGlyThrIleValSerGlyThrLeuLeuAla	87
Db	1210	AGCAGTGTATCAAGTCGCTGTAGTAAAGCCAGGAGATTGATGTTGGGTACTATCTGACA	1259
Qy	88	GlyIleGlyLeuThrSerIleSerGlyProIleGlyIleIleGlyAlaIleIle	107
Db	1270	-----GCCTTGCATCATTTGTATATCA-----GGTGTGGTACTTATA	1308
Qy	108	SerPheGlyThrIleuIleThrValPheThrProAlaGlyGluGlnAsp--LysThrVal	126
Db	1309	TCATTTGAACCTTGCTCCCGCTCTTGGCTGTATCCAGGAGAAAGATCCAAAAAATTT	1368
Qy	127	TrpThrGlnPheIleLysMetGlyGluIlePheValAspThrProLeuThrGluSerIle	146
Db	1359	TGCTCACATTTATGAAACACAGGAGAAACCTTTAAATCAACAATTTCTACAGCTGTA	1428
Qy	147	LysGluLeuLysLeuGlnThrLeuGluGlyPheArgGlnIleLeuGlnSerTyraThr	166
Db	1429	AAAGAAATAGCATTTAGCTCATCTAAATGGTTTTAAAGATGATTAAAGTACGTATGAAGA	1488
Qy	167	AlaLeuAspArgTrpArgLysLeuLysArgLeuGlnAlaProGlyLeuProProSerSer	186
Db	1489	GCATTTAATGATGGAGAGA-----AATCCAAGTCA	1521
Qy	187	AlaLeuGlnGlnAlaAlaLeuThrLeuLysIleArgPheGluAsnValHisAsnAspPhe	206
Db	1522	-----AATACGCCAGATGGTATACACAGAGATTTGAAAGCCATTTTCATTTT	1572
Qy	207	IleArgGluIleProGlyPheGlnLeuGluThrTyrlsThrLeuLeuProIleTy	226
Db	1573	GTAAGCAATATGGCCCACTCCACTTCCTCCAGTATGACATATATTTAAGTTGCAT	1632
Qy	227	AlaGlnAlaAlaAsnPheHisLeuAsnLeuLeuGlnGlnGlyAlaGluLeuAlaAspGlu	246
Db	1633	ACAGAGCTGCAGAAATTTACATTTGATTTATTAACACAGAGTGTACAAATTCGCGAGATA	1692
Qy	247	TrpAsnAlaAspIleHisProSerGlnIleGluProAsnAlaGlyThrSerAspAspTy	266
Db	1693	TGGATATGACATCAACACACATTTACACCATGTGTAAGTACATCAGTACT-----TAT	1743
Qy	267	TyrlsLeuLeuLysGluAsnIleProLysTySerAsnTyrcysAlaAsnThrTyArg	286
Db	1744	TATGACGAGCTATTTGGTATATATTAAGAAAGATATATATTTATTTGCACACAAATACAT	1803
Qy	287	ThrGlyLeuLysAsnLeuAspArgAspGluProAsnMetLysTrpSerIlePheAsnAspTy	306
Db	1804	AAAGGATTTGAATTCACCTTAAGAAATACAGAAAAATACACATGAGATGCTTAATACATAT	1863
Qy	307	ArgArgTyMetCThrIleThrValLeuAspThrIleSerGlnPheSerLeuTyAspIle	326
Db	1864	CGTGAGAAATAGACCTTAATATGATTTGATCTTGGCGCAACTTTCCCTTTTATGATATA	1923
Qy	327	LysArgTyraArgAspSerIleGlyGlyIleGluValLysGlyIleLysAsnGluLeuThr	346
Db	1924	CGCTCTTTT-----CCMAAGAGATAGAACTTGAAATTTTACA	1959
Qy	347	ArgGluIleTyThrThrGluIleAsnPheAspArgLeuProGlnLeuAspValGlnPro	366
Db	1960	AGAGAGGTTTATACAGTTTATAGATCATTTAACACAGACACA-----	2001
Qy	367	AsnLeuAlaThrMetGluTyraAsnLeuThrArgAlaSerPheLysLeuPheSerPheLeu	386
Db	2002	-----GGGCTATTACTTGCTG	2019
Qy	387	GluGlnPheIlePheTyThrGluAsnThrAsnPheGlyAsnArgLeuValGlyIleSer	406
Db	2020	TCAGATATATGATTATTAACACGAGAGGTGTGGACAAAGGCGATATTTATTAACAGATAT--	2076

```

QY AsnArgAspAlaProThrIleTyrSerAsnThr-----IleThrGluThrLeuValGly 422
Db |||:::|||||||:::|||||||:::|||||||:::||||||| 2077 ---CGAGGTCTAAATATATTACTACGGTAATCAATTTTTCAGATGAATAAATATTATAGT 2133
QY GluArgThrGly---SerProThrThrlLeuThrIleArgProPheGlnSerTyrLeuVal 442
Db |||:::|||||||:::|||||||:::|||||||:::||||||| 2134 AATACAAATAGATTAAAGTAGAGCAGCTCATTAATCATATTATACCAGCCGAA---TTTATAGCT 2190
QY SerIleValThrAspArgGlnSerProValSerProIleGlnProHisPheIleIle 462
Db :::::|||||||:::|||||||:::|||||||:::||||||| 2191 CACTTAAGCATAAACCGTCTTTTCAACAACAATGCTGGTATTAATTAAGTTATACAGTTTA 2250
QY AsnGlnIleGluLeuTyrLeuAsnGlySerSerAsnThrLeuTyrTyrSerAlaGly 482
Db |||:::|||||||:::|||||||:::|||||||:::||||||| 2251 ATTCAAAAATTCGATTATCACACTTTTAAAAACATATAT-----GAATATCAAAAAAT 2304
QY GlySerLeuSerAsnTyr-----GlnAsnThrThrPhePheGlnPheProArgLys 499
Db ::::::|||:::|||||||:::|||||||:::||||||| 2305 TTTATATGAAATATCAAAATGAACTCAAGCAAAATCA----- 2343
QY LysAspCyAsnLeuValIleAspProGlyCysSerProAsnPheAsnAsnTyrSerHis 519
Db |||:::|||||||:::|||||||:::|||||||:::||||||| 2344 -----AACTATCTCAATGATTATGAGTGCTTCA--AACAGCCAATAATCAAAAT 2391
QY IleLeuSerHisPheSerLeuPheThrTyrSerTyrValIleGlyLeuGlnIle 539
Db |||:::|||||||:::|||||||:::|||||||:::||||||| 2392 AATTATCTCATTTTCCATTATATCATCCAC-----AACTTAGAGCTT 2433
QY LeuAsp-----ThrGlyValLeuGlyTyrThrHisSerSerValAspArgTyr 555
Db ::|||:::|||||||:::|||||||:::|||||||:::||||||| 2434 GCTGAGATATTTTCACTCATATTTGCATATGAGTTGGACACACATATAGTATACTCCCA 2493
QY AsnAlaIleSerAspLysIleIleThrMetIleProAlaIleLysGlyAsnAsnLeuAsp 575
Db |||:::|||||||:::|||||||:::|||||||:::||||||| 2494 AATTATATATCAGAAGGTGATGATGACAAATCCCATGTGTAAGTTACGTTACGAAGTT-- 2550
QY ThrAsnSerLysValIleGlnGlyProGlyHisThrGlyGlyAsnLeuValTyrLeuGln 595
Db |||:::|||||||:::|||||||:::|||||||:::||||||| 2551 ACTAACATTCAGATTATAGAGGACCAAGCTTTTACAGTGGAGATTATATAGAACTTCGT 2610
QY SerGlnArgLeuGlnIleThrCysGluThrProAsnSerThrGlnSerTyrPheIle 615
Db ::|||:::|||||||:::|||||||:::|||||||:::||||||| 2611 GAT-----AAATGTTCTAATTAATGTAAAGCT--AGTTCTTAAAAAATAAGCTATA 2661
QY ArgLeuArgTyrAlaThrAsnGlyAlaGlyAsnThrLeuProAsnIleSerLeuThrIle 635
Db |||:::|||||||:::|||||||:::|||||||:::||||||| 2662 AGCTATATTTTATGCTGCMAAATAGCAATAGCTGTATCAATAGACGTAGTATCCGA 2721
QY ProGlyValIleGlyIleProProGlnArgLeuAsnAsnThrPheSerGlyThrAsnTyr 655
Db |||:::|||||||:::|||||||:::|||||||:::||||||| 2722 GCAGAGATTCA-----TTGCAACTACCTTTTCTAGAAAAGGAC 2763
QY AsnAsn-----LeuGlnTyrGlyAspPheGlyTyrPheGlnPheProSer 670
Db |||:::|||||||:::|||||||:::|||||||:::||||||| 2764 AATATATTTTCAATTCACAAGACCTTAAGATTATAGATTTCATATTCATACCTTTAGTT 2823
QY ThrValThrLeuProLeuAsnArgAsnIleProPheIlePheAsnArgAlaAsp----- 688
Db ::|||:::|||||||:::|||||||:::|||||||:::||||||| 2824 GATATGTGAATTAACCGAAAGTGAAGAAATTCATATTCATTGAAACGACAGAGATATAT 2883
QY ValSerAsnSerIleLeuIleIleAspLysIleGluPheIleProIleThrSerSerMet 708
Db |||:::|||||||:::|||||||:::|||||||:::||||||| 2884 GAGAGGAGAGATCTTTTATATGATTAATATAGATTCAAACTATATGATGAATAATAT 2943
QY HisGlnAsnArgGluTyrGlnLysLeuGlnThrIleGlnThrLysIleAsnThrPhePhe 728
Db ::|||:::|||||||:::|||||||:::|||||||:::||||||| 2944 -----ACTAATGAATGAATTTTGAAGAGGCAAAAGCAAGCATGTATATATTT 2994
QY ThrAsnHisThrLysThrLeu 725
Db |||:::|||||||:::|||||||:::|||||||:::||||||| 2995 ATTAACGACAAACGCTTTG 3015

```

Search completed: January 20, 2006, 17:46:31  
Job time : 1177 secs

---

GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - nucleic search, using frame p2n model

Run on: January 20, 2006, 16:31:25 ; Search time 9649 Seconds  
(without alignments)  
4329.976 Million cell updates/sec

Title: US-10-783-417-2  
Perfect score: 3869  
Sequence: 1 MNQNNNDNNEYEIDSHSTSPY.....KLETIQTKNPFTNHTKTL 735

Scoring table:  
BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 5883141 seqs, 28421725653 residues  
Total number of hits satisfying chosen parameters: 11766282

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL=frame+ p2n.model -DEV=x1h  
-O=/cgn2.1/USPTO.spool/US10783411/runat\_20012006\_095010\_23867/app\_query.fasta\_1.903  
-DB=GenEmbl -QFMT=fastap -SUFFIX=p2n.rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=blomsun62 -TRANS=human40.cdi -LIST=45  
-DOALIGN=200 -THR\_SCORE=DCT -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptco -NORM=ext -HEAPSIZE=500 -MILEN=0 -MAXLEN=200000000  
-USER=US10783411@cgn2.1 6416 @runat\_20012006\_095010\_23867 -NCPU=6 -ICPU=3  
-NO\_WMAP -NRC SCORES=0 -WAIT -DSPBLOCK=100 -DEV TIMEOUT=120 -WARN TIMEOUT=30  
-THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7 -YGAPOP=10  
-YGAPEXT=0.5 -DELOP=6 -DELEXT=7

## Database :

GenEmbl:.\*  
1: gb\_ba:.\*  
2: gb\_in:.\*  
3: gb\_env:.\*  
4: gb\_ov:.\*  
5: gb\_ov:.\*  
6: gb\_pac:.\*  
7: gb\_ph:.\*  
8: gb\_pr:.\*  
9: gb\_ro:.\*  
10: gb\_sts:.\*  
11: gb\_sy:.\*  
12: gb\_un:.\*  
13: gb\_vl:.\*  
14: gb\_hcg:.\*  
15: gb\_pl:.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

Result No.	Score	Query Match	Length	DB ID	Description
1	3869	100.0	2208	6	CQ868305 Sequence
2	3314	85.7	2235	6	CQ868307 Sequence
3	3104.5	80.2	2085	6	CQ868309 Sequence

## SUMMARIES

4	1061.5	27.4	3543	1	BTTOXD1	Y00423 Bacillus th
5	1061.5	27.4	4451	6	I08083	I08083 Sequence 1
6	1061.5	27.4	4451	6	I09103	I09103 Sequence 1
7	1060.5	27.4	4253	1	BACIERH4	D00248 Bacillus th
8	1060.5	27.4	4934	6	E01676	E01676 DNA sequenc
9	1060.5	27.4	127923	1	BTBPTOXIS	AL731825 Bacillus
10	897.5	23.2	3753	1	BACMSQB	M12662 B.churingie
11	897.5	23.2	127923	1	BTBPTOXIS	AL731825 Bacillus
12	880	22.7	3756	6	E00614	E00614 DNA encodin
13	862.5	22.3	4959	1	AP285775	AP285775 Bacillus
14	850	22.0	6009	1	BTB251977	AJ251977 Bacillus
15	849	21.9	4896	1	AF132928	AF132928 Bacillus
16	808.5	20.9	3752	1	AB125059	AB125059 Bacillus
17	784	20.3	3644	1	BTB251978	AJ251978 Bacillus
18	775.5	20.0	3507	1	BTB04365	U04365 Bacillus th
19	775.5	20.0	3507	6	I25972	I25972 Sequence 3
20	759	19.6	3684	1	BTTOXD2	X07423 Bacillus th
21	759	19.6	3684	6	I08884	I08884 Sequence 2
22	757.5	19.6	3535	6	I06096	I06096 Sequence 2
23	754	19.5	3684	1	BTTOX	X07082 Bacillus th
24	752.5	19.4	4186	1	BACISRH3	D00247 Bacillus th
25	752.5	19.4	4186	6	E01905	E01905 genomic DNA
26	746	19.3	3642	1	AB193814	AB193814 Bacillus
27	709	18.3	3668	1	D88381	D88381 Bacillus th
28	705	18.2	4391	1	BTJEG65	Y07603 B.churingie
29	705	18.2	6698	1	AB116651	AB116651 Bacillus
30	703	18.2	3633	6	CS130948	CS130948 Sequence
31	703	18.2	3633	6	AK543926	AK543926 Sequence
32	703	18.2	6613	6	CS130963	CS130963 Sequence
33	703	18.2	6613	6	AK543951	AK543951 Sequence
34	696.5	18.0	4056	1	BACCRXD2	M20242 B.churingie
35	696	18.0	3536	6	E01029	E01029 DNA sequenc
36	689.5	17.8	4188	6	BD177494	BD177494 Polypepti
37	689.5	17.8	4188	6	AK670919	AK670919 Sequence
38	687	17.8	3931	1	AB161456	AB161456 Bacillus
39	678.5	17.5	2003	6	AK543930	AK543930 Sequence
40	678.5	17.5	2022	6	CS130958	CS130958 Sequence
41	678.5	17.5	2022	6	AK543940	AK543940 Sequence
42	677	17.5	8572	1	AB116649	AB116649 Bacillus
43	677	17.5	8605	1	AB116650	AB116650 Bacillus
44	676	17.5	1953	6	I25973	I25973 Sequence 5
45	676	17.5	1977	1	BTCADE	X17123 Bacillus th

## ALIGNMENTS

RESULT 1	CQ868305	2208 bp	DNA	linear	PAT 13-SBP-2004
LOCUS	CQ868305	Sequence 6 from Patent WO2004074462.			
DEFINITION	CQ868305				
ACCESSION	CQ868305.1	GI:51998351			
VERSION					
KEYWORDS					
SOURCE					
ORGANISM	Bacillus thuringiensis				
	Bacillus thuringiensis				
	Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus; Bacillus				
	cereus group.				
REFERENCE	1				
AUTHORS	Carozzi, N., Hargiss, T., Kozief, M.G., Duck, N.B. and Carr, B.				
TITLE	Delta-endotoxin genes and methods for their use				
JOURNAL	Patent: WO 2004074462-A 6 02-SBP-2004;				
	Athenix Corporation (US)				
FEATURES					
source	1..2208				
	Location/Qualifiers				
	/organism="Bacillus thuringiensis"				
	/mol_type="unassigned DNA"				
	/db_xref="taxon:1428"				
	1..2208				
	/note="unnamed protein product"				
	/codon_start=1				
	/transl_table=11				
	/protein_id="CAH33948.1"				
	/db_xref="GI:51998352"				
CDS					



|||||  
Db 1801 GAAATTACATGGAACCTCTAATCTACACAACTTATTCTAGACTCGATATGCT 1860  
QY 621 ThrAenGlyValAGlySerThrLeuProAenIleSerLeuThrIleProGlyValIleGly 640  
Db 1861 ACAAATGGTGGGAAATACTCTCTTAATATATCTCTTACATACAGAGAAATGGA 1920  
QY 641 IleProGlnIleArgLeuAenAenThrPheSerGlyThrAenTyraAenLeuGlnTy 660  
Db 1921 ATRACCTCTACAGACTCAACAACTTTTCTGGTACAAATTAATATATTTACATAC 1980  
QY 661 GlyAspPheGlyTyThrPheGlnPheProSerThrValThrLeuProLeuAenArgAenIle 680  
Db 1981 GGAGATTTTGGGATTTCCATTTCCAGATGACATTAACCTTAAATGGAACATA 2040  
QY 681 ProPheIlePheAenAenArgAlaAspValSerAenSerIleLeuIleIleAspIleGly 700  
Db 2041 CCATTTATATTAATATGTCAGATGATCAAAATTCATTTATCATTAATAAATTTGAA 2100  
QY 701 PheIleProIleThrSerSerMetHISGlnAenAenGlyLeuGlnIleGlyLeuThrIle 720  
Db 2101 TTTATACCAATTAATCTCTCTATGACCAAAATAGGAAACAAATAATTAAGAACTATC 2160  
QY 721 GlnThrIleIleAenThrPhePheThrAenHISThrIleThrLeu 735  
Db 2161 CAAACAAATTAATATCATTTTCAAAATCATACAAACAACTT 2205

RESULT 2  
CQ868307 2235 bp DNA linear PAT 13-SEP-2004  
LOCUS Sequence 8 from Patent WO2004074462.  
DEFINITION CQ868307  
ACCESSION CQ868307  
VERSION CQ868307.1 GI:51998353  
KEYWORDS

SOURCE Bacillus thuringiensis  
ORGANISM Bacillus thuringiensis  
Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus; Bacillus cereus group.

REFERENCE 1  
AUTHORS Carozzi, N., Hargiss, T., Koziel, M.G., Duck, N.B. and Carr, B.  
TITLE Delta-endotoxin genes and methods for their use  
JOURNAL Patent: WO 2004074462-A 8 02-SEP-2004;  
Athenix Corporation (US)

FEATURES  
source Location/Qualifiers

CDS  
1..2235  
/note="unnamed protein product"  
/codon\_start=1  
/transl\_table=11  
/protein\_id="CAH33949.1"  
/db\_xref="GI:51998354"  
/translation="MNQNNNNEVEIILDSKNLSVPSNRNIDHSRYPTNNPNQNTN  
YKEMLMCGNTQYQDNERTPASADTIAAASAGTYSGLKAGIGLISISEPFIGIG  
AIIISFGTITIVFWPAGBODKVTWQIFVDTPTLIESIKQIKDQTLSEFROI  
QSYNTALDWRKRLKQAPGLPSSALQAPALTLKIRFNVNDPRLREIRPQLEPKY  
TLPLPYAQDAANFHLNLQCGABLADEBNADHPISQIBENAGISDYYLKLKENIPKY  
SNVCANTYREGINKLNEBNPMSIENDRYRWITIVLDTIAQSFYDIKRYDSIGR  
IGIKTELRLREIYTRINEINFDRLTYLEIOPNALIMEYNLRSGLRLESPIDELIFTRKN  
ERYGNRLVGIANRNSTYATTGTEIIEYGRTPPTTKTILPPESYKVSIVTBRQVPT  
SPPTNYFTIINQIELYNLSPNSKLTYSAGNLSNDKTTDQFQPKQCKPIINENC  
LPSVNSYSHIISQFLFNSYKIGLANLTYGALGWTSSVNRUNAISDKIITMIPA  
IKGNSLDTNSKVLIEGHTGNNLVYQSGRLIECTRTNSQSYIRLRYATNGAGN  
TLPNISLTIIPGVIQIPORLNTFSGTNNNOYDGFQFQPSPTVTLPLNNIPITF  
NRADVSNLIIIDKIEFIPITSSVRNKRQKLEFIIQTINIFPNTHTNNTINIAATN  
YIID"

ORIGIN

Alignment Scores:

Pred. No.: 5..39e-241 Length: 2235  
Score: 3314.00 Matches: 638

Percent Similarity: 91.34  
Best Local Similarity: 86.94  
Query Match: 85.74  
DB: 6  
Gaps: 3

US-10-783-417-2 (1-735) x CQ868307 (1-2235)

QY 1 MetAenGlnAenAenAenAenAenGluTyThrGluIleIleAenSerHISThrSerProTy 20  
Db 1 GTGAATCAAAATTAAT---AATTAATGAATATGATTAATCGATTCAAGAATTAATCTTAT 57  
QY 21 PheProAenAenAenAenAenAenAenAenAenAenAenAenAenAenAenAenAenAen 40  
Db 58 CCTTCAACAGAAATATGATGATCATTCGATACCTTACACAAATATCAATCAACCA 117  
QY 41 LeuGlnAenThrAenTyThrIleGlyLeuThrLeuAenMetCysGlnGlyAenThrGlnTy 60  
Db 118 TTACAAACACAAATTAACAAAGATGCTCAATATATGTCAGAGGAAATACAAATATGCT 177  
QY 61 AspAenPheGluThrPheAlaSerAlaAspThrIleAlaIleValSerAlaGlyThrIle 80  
Db 178 GATTAATTCGACACATTTGCTATGCTGATACAAATTCGTCGATTAATGTCAGAGTACTAT 237  
QY 81 ValSerGlyThrLeuLeuAlaGlyIleGlyLeuThrSerIleSerGlyProIleGly 100  
Db 238 GTATCCGACTCTGTTAGCCGTTATAGTGGCTCACCTTATATCCGACCGAATGGA 297  
QY 101 IleIleGlyAlaIleIleIleSerPheGlyThrLeuIleThrValPheTrpProAlaGly 120  
Db 298 ATTAATAGGTCATATATATATATCTTTGGTACCTTAATCACTGCTTTTGGCCCGGGA 357  
QY 121 GluGlnAenAenAenAenAenAenAenAenAenAenAenAenAenAenAenAenAenAen 140  
Db 358 GACACAGACAAACAGATGACACAAATTTATTAATAGGGAATTTTGTGATACA 417  
QY 141 ProLeuThrGluSerIleGlyGlnLeuAenGlnThrLeuGlnGlyPheArgGlnIle 160  
Db 418 CCGTTACAGAAAGCATTAACAGCTAAAGTTACAACTTTGAAAGATTTAGCAAAATA 477  
QY 161 LeuGlnSerTyThrAenThrAlaLeuAenAenAenAenAenAenAenAenAenAenAen 180  
Db 478 TTACAAAGCTTAATTAACAGCATTAATGATGTCGAAATAATTAATAAGCATCAAGCTCT 537  
QY 181 GlyLeuProProSerSerAlaLeuGlnGlnAlaAlaLeuThrLeuIleIleArgPheGly 200  
Db 538 GGATTAACCAACCATTCATGACATTAACACAGCTGCTTGAATCTTAATAATGATTTGAG 597  
QY 201 AenValHISaenAenPheIleArgGluIleProGlyPheGlnLeuGluThrTyThrTyThr 220  
Db 598 AATGTTCAACATGATTTTATTCGAGAAATCCTGCTTCCAACTTGAACTTATTAAGAG 657  
QY 221 LeuLeuLeuProIleTyThrAlaGlnAlaAlaAenPheHISLeuAenLeuGlnGlnGly 240  
Db 658 CTATTACTACTATTTATGCGGACGCTGTAATTTTCAATTAATTAATTAACAAAGCT 717  
QY 241 AlaGluLeuAlaAspGluThrAenAlaAspIleHISProSerGlnIleGluProAenAla 260  
Db 718 GCTGAATGGCTGAAGAATGAATGACATTAACCTTCAAAATGAACCTATAGCT 777  
QY 261 GlyThrSerAenAenAenAenAenAenAenAenAenAenAenAenAenAenAenAenAenAen 280  
Db 778 GGAACATCGATGATTAATTAATCTTTTAAGAAATATATCTTAATATATGTAATCTAT 837  
QY 281 CysAlaAenThrTyThrGlyThrGlyLeuAenAenAenAenAenAenAenAenAenAenAen 300  
Db 838 TGTGCAAAATCACTATAGGAAGACCTAAATAAATTCGAAAGCAACTAATATGAGATG 897  
QY 301 SerIlePheAenAenAenAenAenAenAenAenAenAenAenAenAenAenAenAenAenAen 320  
Db 898 AGTATATTAATGATTAATGAGATTAATGATTAATGATTAATGATTAATGATTAATGATTA 957  
QY 321 PheSerLeuTyAspIleTyThrGlyThrGlyAenSerIleGlyGlyIleGluValIleGly 340

Db	958	TTTTCTTTTATGATATATAAGAGATACAAAGTTCAATATAGCA-----AGATATAGTGGC	1011
Qy	341	ILelysaengGluLeuThrArgGluIleYThrThrGluIleasnPhesapArgLeuPro	360
Db	1012	ATTTAAATCGAATCTTACCAAGAAATTTATATCACTGAAATTAATTTTATGACCGTCTTACT	1072
Qy	361	GlnLeuAsgValGlnProAsnLeuAlaThrMetGluYrAsnLeuThrArgAlaSerPhe	380
Db	1072	TACCTTGAATTCACCAATCCCAATTCGCTATATAGAAATTAATTTTACACGCTCAGGGCTT	1132
Qy	381	LysLeuPheSerPheLeuGlnGlnPheIlePheYrThrGluAsnThrAsnPhesGluYasn	400
Db	1132	AGATTATTTTCATTTTATGATGATCACTTATATTTTATACAAAAAAGAAACGTAACGGGAT	1192
Qy	401	ArgLeuValGlyIleSerAsnAspAlaProIleYrSerAsnThrIleThrGluThr	420
Db	1192	CGTTAGATGGTATGGCATGGATGGTATAGATCTACTTATGGTACACAGAACTGAAATTT	1252
Qy	421	LeuYrGlyIleArgThrGlySerProThrIleYrThrIleArgProPhaGlnSerYr	440
Db	1252	ATATATGAGAAAGAAACAGGTCCACCCCAACAAAACTTTATATCAATTTGAAATCCAT	1312
Qy	441	LysValSerIleValThrAspArgIleSerProValSerProIle---GlnProHis	459
Db	1312	AAAGTTTCATTTGATGATGATGACATGACAAAGTAATCTTACTTCCCTTTCTTACATATAC	1372
Qy	460	PheIleIleasnGlnIleGluLeuYrLeuAsnGlySerSerAsnAsnThrLeuYrYr	479
Db	1372	TTTACATATTAATCAATTAATGAATCTTATTTAATTAATTCACCTAGTAATTAATTAATCAATAT	1432
Qy	480	SerIleGlyIleYrSerLeuSerSerYrGlnAsnThrPhePheGlnPheProArgYs	499
Db	1432	TCAGCTGGGGGAAATTTATCTATATATATATAAAAAAACACTGATTTTCAATTTCCGTATAAA	1492
Qy	500	LysAspCysAsnLeuValIleAspProGlyCysSerProAsnPhesAsnIleYrSerHis	519
Db	1492	AAAGCTGTAACCAATTTATATCAAAATCTTTACCAAGCTATATATAGTATATGCTAT	1552
Qy	520	IleLeuSerHisPheSerLeuPheThrYrSerYrValIleGlyLeuGlnLeuGlnIle	539
Db	1552	ATTTATATCCCAAGTTTCTTTATTTATATATTTCTTATTAATAATGATGACGCTAAATATATA	1612
Qy	540	LeuAspThrGlyValLeuGlyYrThrHisSerSerValAspArgYrAsnAlaIleSer	559
Db	1612	TTATATACAGGTGACATTAGAGACACACACAGTACTGTAATAGAAATATGACATATACA	1672
Qy	560	AspYrsIleIleThrMetIleProAlaIleIleYrGlyYasnAsnLeuAspThrAsnSerYs	579
Db	1672	GATTAATAATATATACATGATCCCAAGCAATCAAAAGGTAAACGCTTGATACAAACTCTTAAG	1732
Qy	580	ValIleGluGlyProGlyYHisThrGlyYasnLeuValYrLeuGlnSerGlnIleYrG	599
Db	1732	GTAATTTGAAGAGACCTGGTCAATACAGAGAAACTGGTTATTTTACAAAGTCAAGGGGT	1792
Qy	600	LeuGluIleThrCysGluThrProAsnSerThrGlnSerYrPheIleArgLeuArgYr	619
Db	1792	TTAGAGATTACATGTATGAACCTCTATTTCTACACATCTTATATCATTTAGACTTGATAC	1852
Qy	620	AlaThrAsnGlyValGlyAsnThrLeuProAsnIleSerLeuThrIleProGlyValIle	639
Db	1852	GCTACAAATGGCGTGAATAATCTCTTCTAATATATCTTTACAAATCCAGAGATATATA	1912
Qy	640	GlyIleProProGlnArgLeuAsnAsnThrPheSerGlyYrAsnYrAsnAsnLeuGln	659
Db	1912	GGAATATCACACCTCAACGACTCAACAACCTTTTCTGGTACAAATTAATATATATTTACAA	1972
Qy	660	TyrGlyAspPheGlyYrPheGlnPheProSerThrValThrLeuProLeuAsnArgAsn	679
Db	1972	TACGAGAGATTTGGGTATTTTCCAAATTTCCAAAGTACAGTAACTTAAATGCAAAAC	2032
Qy	680	IleProPheIlePheAsnArgAlaAspValSerAsnSerIleLeuIleIleAspYrsIle	699
Db	2032	ATACCATTTATATTTAATCGTGCAATGATATCAATTTCAATTTTATATCATTTGATTTAAATTT	2092

QY	700	GlubhellePcoIleThrSerSerMechiSegIlnaSenAagGluGlyGlnIlyLeuGluIuThr	719
Db	2092	GAATTATTAACCAATTACTCTCTGTACGCCAAAATAGAGAAAAACAATAATTGAAACT	2157
QY	720	IlleginThyIleAenThrPhePheThraSenHisThrIlys	733
Db	2152	ATCCAAACAAAATTAATACATTTTTCACAAATCATACAAA	2193
RESULT 3			
LOCUS	CQ868309	2085 bp	DNA
DEFINITION	Sequence 10 from Patent WO2004074462.	linear	PAT 13-SEP-2004
ACCESSION	CQ868309		
VERSION	CQ868309.1	GI:51998355	
KEYWORDS			
SOURCE	Bacillus thuringiensis		
ORGANISM	Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus; Bacillus cereus group.		
REFERENCE	1		
AUTHORS	Carozzi, N., Hargiss, T., Koziel, M. G., Duck, N. B. and Carr, B.		
TITLE	Delta-endotoxin genes and methods for their use		
JOURNAL	Patent: WO 2004074462-A 10 02-SEP-2004;		
FEATURES	Athenix Corporation (US)		
source	Location/Qualifiers		
	1..2085		
	/organism="Bacillus thuringiensis"		
	/mol_type="unassigned DNA"		
	/db_xref="taxon:1428"		
	1..2085		
	/note="unamed protein product"		
	/codon_start=1		
	/transl_table=1		
	/protein_id="CAH33950.1"		
	/db_xref="GI:51998356"		
	/translation="MCGQNTQYGDNPETFPADDTIAVSAQTVSGTILADIGLSTIS		
	SGRIPIGAILIISFTLITVFWPAGSDQKWTQPIKGEIFVDTPLTSLKQLQ		
	LESGRQILQSYNTALDDWRKLRQLQAPLPPSALQQALTLKRPENVAHDFIRE		
	GFQLEIYKTKLLPIYQAQAAFHNLMLQQGALADENWADHIPSOIEPVAGTSDDYK		
	LKENIPKXSNCAVTYREGLNKLNENPMEKSIENDYRWYITVLDITADFSPYDI		
	LYKDSIRIGGIKTELREIYTTREIYTTINDRLLYELIOPLAIMEYNLRSGLRFSPL		
	ELIYTTKREIYGNRLVGIANRNSRYATYTRIRIYGERSTPPTKTLIPESYKYST		
	TDROVTPISPPNPVFTINQLELYLNNSPSKRLTYSAQGNISNDKTTDPQFPVKDD		
	KPIINMPLPSYNSYSHLSQFSLFNSYKGLALNITLYTALQMTSSVRNNAISIS		
	KIITMPLAIKENSVDLNSKVLREGFGHNGNLVLYOSGRLEITCRTPSTOSYIRLL		
	YATNGAGNTLPIISLITPGVIGIPROSLNMFSGTNNINLOYDGPVGFPSVTLPI		
	NRNIPPIFNRAVDVNSILIIDIKIEFIPITTSVRNRKQLKETTQTKINFTNHTTK		
	TLNTEATNYDID"		
ORIGIN			
Alignment Scores:			
Pred. No.:	3.51e-225	Length:	2085
Score:	3104.50	Matches:	598
Percent Similarity:	91.9%	Conservative:	30
Best Local Similarity:	87.6%	Mismatches:	52
Query Match:	80.2%	Indels:	3
DB:	6	Gaps:	2
US-10-783-417-2 (1-735) x CQ868309 (1-2085)			
QY	52	MetCysgInglYaeThrGlnTyrcIYAspaanPhegluThrPheAlSeraIaapThr	71
Db	1	ATGTCACAGGAATACACAAATATGTCATTAATTCGAGACATTTGCTAGTGCATACA	60
QY	72	IlleaIaValSerAlaGlyThrIleValSerGlyThrLeuLeuAlaGlyTlleglyI	91
Db	61	ATTGCTGAGTAGTAGCAGAGTACTATGTATCCGGTACTCTGTTAGCCGATATAGTGG	120
QY	92	LeuThrSerIleSerGlyProIleGlyIleIlegIlaIleIleIleSerPheGlyThr	111
Db	121	CTCCTCTTATATCCGACCGAATAGGAATATATAGTGTCTATTAATATATCTTTGACC	180

QY 112 LeuIleThrValPheTrpProAlaGlyGlnAspLeuThrValTrpThrGlnPheIle 131  
 DB 181 CTRATACCTGCTTTTGGCCCGGGAGAAACAAGACAGTATGACACATTTATTT 240  
 QY 132 LysMetGlyGlnIlePheValAspThrProLeuThrGlnSerIleLysGlnLeuLysLeu 151  
 DB 241 AAAATGGGAGAAATTTTGTGTATACCGTTAAACGAAAGCATAAACGCTAAAGTTA 300  
 QY 152 GlnThrLeuGlnGlyPheArgGlnIleLeuGlnSerTrpAsnThrAlaLeuAspAspTrp 171  
 DB 301 CAAACTTATGAAAGATTATGACAAATATTAACAAAGCTATATATACACATTAGATGATG 360  
 QY 172 ArgLysLeuLysArgLeuGlnAlaProGlyLeuProProSerSerAlaLeuGlnGlnAla 191  
 DB 361 AGAAATTTAAAGATACACAGCTCTGATTTACCAATCATACACATTTACACAGCT 420  
 QY 192 AlaLeuThrLeuLysIleArgPheGluAsnValHisAsnAspPheIleArgGlnIlePro 211  
 DB 421 GCCTGACTCTTAAATACGATTGGAATGTTCAACATGATTTTATTCGAGAAATACCT 480  
 QY 212 GlyPheGlnLeuGlnIleThrTrpLysThrIleLeuLeuProIleTrpAlaGlnAlaAsn 231  
 DB 481 GGTTCCTCAACTGAACTTATATAAAGCTATATTACTATATTTATGCGAAGCTGCTAAT 540  
 QY 232 PheHisLeuAsnLeuLeuGlnGlnGlyAlaGluLeuAlaAspGluTrpAsnAlaAspIle 251  
 DB 541 TTTTCATTTAAATTTATTTACACAGAGTCTGATTTGGCTGATGAAAGGATGCAATTA 600  
 QY 252 HisProSerGlnIleGluProAsnAlaGlyTrpSerAspAspTrpTrpLysLeuLys 271  
 DB 601 CATCTCTCAAAATTTGACCTATATGCTGGAACATGACATGATATATTAATTTAA 660  
 QY 272 GluAsnIleProLysTrpSerAsnTrpCybAlaAsnThrTrpArgThrGlyLeuLysAsn 291  
 DB 661 GAAAAATATCTTAAATATAGTAACTATGCGAAATACCTATAGAAAGACCTAAATTA 720  
 QY 292 LeuArgAspGluProAsnMetLysTrpSerIlePheAsnAspTrpArgTrpMetThr 311  
 DB 721 CTTGGAAGCAAGACCTATATGAGATGAGATATTTATGATTTATGCAAGATATGACT 780  
 QY 312 IleThrValLeuAspThrIleSerGlnPheSerLeuTrpAspIleLysArgTrpArgAsp 331  
 DB 781 ATTACTGATTTAGATACCTATCGCTCAATTTCTTTTATGATATTAAGATACAAAGAT 840  
 QY 332 SerIleGlyGlyIleGluValLysGlyIleLysAsnGluLeuThrArgGluIleTrp 351  
 DB 841 TCATATGGA-----AGAAATAGGTGCAATTAAACGAACTTACAGAGAAATTTATACA 894  
 QY 352 ThrGluIleAsnPheAspArgLeuProGlnLeuArgValGlnProAsnLeuAlaThrMet 371  
 DB 895 ACTGAATATTAATTTGACCGTCTTACTTACCTTGAAATTCACCCATCTCGCTATATG 954  
 QY 372 GluTrpAsnLeuThrArgAlaSerPheLysLeuPheSerPheLeuGlnPheIlePhe 391  
 DB 955 GAATATTAATTTTACAGTTCAAGGCTTAGATTATTTTCAATTTTATGATGACCTATATTT 1014  
 QY 392 TrpThrGlnAsnThrAsnPheGlyAsnArgLeuValGlyIleSerAsnArgAspAlaPro 411  
 DB 1015 TATACAAAAAATGAAACGTAAGGAAATCGTTTATGTTAGTATGCGAATCGTATATGATCT 1074  
 QY 412 ThrTrpSerAsnThrIleThrGlnThrLeuTrpGlyGlnArgThrGlySerProThrTrp 431  
 DB 1075 ACTTATGCTACGACAGAACTGAAATTTATATATGAGAAAGACAGCTCCACACACACA 1134  
 QY 432 LysThrIleArgProPheGlnSerTrpLysValSerIleValThrAspArgGlnSerPro 451  
 DB 1135 AAAACTTTAATACATTTGATCTTATTAAGTTTCAATGCTATGATGACAAAGTAATCT 1194  
 QY 452 ProValSerProIle--GlnProHisPheIleIleAsnGlnIleGluLeuTrpLeuAsn 470  
 DB 1195 CCTACTCTCCCTTTCTTAAACATATACCTTTACAAATTAATCAATTTATTTTAAAT 1254  
 QY 471 GlySerSerAsnAsnThrLeuLysTrpSerAlaGlyGlySerLeuSerAsnTrpGlnAsn 490

DB 1255 AATTCACCTAGTAATTAATTAACATATTCAGCTGGGGGAAATTTTATCTAATGATAAAAA 1314  
 QY 491 ThrThrPhePheGlnPheProArgLysLysAspCybAsnLeuValIleAspProGlyCyb 510  
 DB 1315 ACNACTGATTTTCAATTTTCCGTATTAATAAAGACTATAACCAATATTAATCAAAATGCT 1374  
 QY 511 SerProAsnPheAsnTrpSerHisIleLeuSerHisPheSerLeuPheThrTrpSer 530  
 DB 1375 TTACCAAGCTATTAATGATTTATGCTATTTTATCCAGTTTCTTTATTTATTAATTTCC 1434  
 QY 531 TrpValIleGlyLeuGlnLeuGlnIleLeuAspThrGlyValLeuGlyTrpThrHisSer 550  
 DB 1435 TATAAATTTGATTTGCGCTTAATATATATATATATACAGGTGATTAAGATGACACACAGCT 1494  
 QY 551 SerValAspArgTrpAsnAlaIleSerAspLysIleIleThrMetIleProAlaIleLys 570  
 DB 1495 AGCTTTAATGAAATATGCAATATTCAGATTAATAATTAATCAATATCCACCAATCCAA 1554  
 QY 571 GlyAsnAsnLeuAspThrAsnSerLysValIleGluGlyProGlyHisThrGlyLysAsn 590  
 DB 1555 GGTATACGCTTGAATACAACTTAAAGTATTTGAAGACCTGCTCATACAGAGAGAAAC 1614  
 QY 591 LeuValIleLeuGlnSerGlnGlyArgLeuGlnIleThrCybGluThrProAsnSerThr 610  
 DB 1615 TTGGTTTATTTTCAAAAGTCAAGGCGTTTATGAGATTTACATGTAAGAACTCTAATTTTACA 1674  
 QY 611 GlnSerTrpPheIleArgLeuArgTrpAlaThrAsnGlyAlaGlyAsnThrLeuProAsn 630  
 DB 1675 CATCTTTATTAATTAATTAAGCTTGTGATACGCTACAAATGCTGGAATATCTTCTTAAT 1734  
 QY 631 IleSerLeuThrIleProGlyValIleGlyIleProProGlnArgLeuAsnAsnThrPhe 650  
 DB 1735 ATATCTTTACATACCGAGATTAATGGAATACCACTCAACGACTCAACACACACTTT 1794  
 QY 651 SerGlyTrpAsnTrpAsnLeuGlnTrpGlyAspPheGlyTrpPheGlnPheProSer 670  
 DB 1795 TCTGTATCAAAATTAATATATATTTACATACGAGATTTTGGTATTTTCCAAATTTCCAACT 1854  
 QY 671 ThrValThrLeuProLeuAsnArgAsnIleProPheIlePheAsnArgAlaAspValSer 690  
 DB 1855 ACAGTAACATTAACCTTAAATGAAACATACCAATTAATTAATCGTGCAATGATATCA 1914  
 QY 691 AsnSerIleLeuIleIleAspLysIleGluPheIleProIleThrSerSerMetHisGln 710  
 DB 1915 AATTCAAATTTAATATGATTAATAATGAAATTAATCAATTAATCTCTGTACGCCAA 1974  
 QY 711 AsnArgGluLysGlnLeuLeuGlnIleThrIleGlnTrpLysIleAsnThrPhePheThrAsn 730  
 DB 1975 ATATGAGAAAAACAAATTAATGAAACTATCCAAACAAATTAATTAATTTTCCAAAT 2034  
 QY 731 HisThrLys 733  
 DB 2035 CATACAAA 2043  
 RESULT 4  
 BTTOX1 3543 bp DNA linear BCT 18-APR-2005  
 LOCUS Bacillus thuringiensis gene for 130 kDa delta-endotoxin.  
 DEFINITION Y00423  
 ACCESSION Y00423.1 GI:40351  
 VERSION delta-endotoxin; endotoxin.  
 KEYWORDS Bacillus thuringiensis  
 SOURCE Bacillus thuringiensis  
 ORGANISM Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus; Bacillus cereus group.  
 REFERENCE 1 (bases 1 to 3543)  
 AUTHORS Ward, E.S. and Bilar, D.J.  
 TITLE Nucleotide sequence of a Bacillus thuringiensis var. israelensis gene encoding a 130 kDa delta-endotoxin  
 JOURNAL Nucleic Acids Res. 15 (17), 7195 (1987)  
 PUBMED 2821500  
 REFERENCE 2 (bases 1 to 3543)

AUTHORS Ward, E.S.  
TITLE Direct Submission  
JOURNAL Submitted (11-MAY-1987) Ward E.S., University of Cambridge,  
Department of Biochemistry, University of Cambridge, Tennis Court  
Rd., Cambridge CB2 1QM  
COMMENT \*strain: var:israelensis;  
Data kindly reviewed (12-JAN-1988) by Ward E.S.  
FEATURES  
source  
1. 3543  
/organism="Bacillus thuringiensis"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:1428"  
1. 3543  
/note="unnamed protein product; endotoxin (AA 1-1180)"  
/codon\_start=1  
/transl\_table=11  
/protein\_id="CAA68485.1"  
/db\_xref="GI:40352"  
/db\_xref="GOA:P16480"  
/db\_xref="UniProt/Swiss-Prot:P16480"  
/translation="MNPKYONKNEYETLNASQKLNISNNYTRYPISNSPKQLQSTNY  
KDLNLMCOONOYGDPEFTIDSGELSAVTIVGKTLFGFTPELAGLIGFTLIP  
VLFPADQSNMTSDEITOTKNTIKKEIASITVSNANKLINSFNVTISVHHLKWTEN  
NPNPQNTDVRTOQLVHYHONVYBELVNGCPNPSQCDYNTILVLSYQAAHL  
TVLNDVAFKAVLKNRPDIPLAIDIVYVTKALIEDTNTCVTVYKGLMLIK  
TTPSNLDGNNMNTYNTRTKMTAVLDVALFPNVGVKYPGIVQSELREIYVL  
NFESPKYDYFOYQEDSLTRPHLFTLDSLNFYKQKTPNPFTHSHMFFYLLD  
NISQSVFNGHNVTDLKSLGLATNIVPLNTLSNKLNDYNNISKDPFITNG  
TRLKEKLJASGGITVDVKNRIKGLPIKRENGNGLTPEYDNNVSHLISFKIS  
IPATKYOTVTPAHTSSVDPKNTTYHTLTOIPAVKANSJGTSKTYVQGHKQGLD  
IDPKHFKITCOHSPQOSYFRIRIKASNGSNTAVNLISI PGABGLMLNPFSG  
TDYTNLKTKDYLEFSNEKFPAPQNI SLVFNRSVDYNTVLIIDKLEPIRISIR  
EDREKQKLETVQOILNTEYANPIKNTLSDNTTIOEDPIFKGHYLNHSGARDIL  
LDEVENAKOLESRRVNLONGDEBSATLGMTSDNTTIOEDPIFKGHYLNHSGARDIL  
GTFPFTYFORKIDESKLPYTRIVYRGVSGKQVLEVSRYGEEDLIMVPADLNY  
LYSPFDEGSRCTSAVPANIGTSMKSCQVDRKRYVCKMSQSPFTIDTGA  
LDTNENIGWVWFKISPDGASLDNLEVISBGPIDGALSRYKMEKRMQDQAKR  
SETQQAIVVAQALDALFTNVDALQDPTLAQIQVAEYVQSI PYYNWMLSVPG  
NMYDIYVELDARVAQALYDINIIKNGDFTQGMHGNADVOQIDISVYLS  
NMSAGVSONVHLQHNHGYLVGLIAKKEGPGNGYVLMMEENQELFTSDEGYTK  
TYDVPDTRVRIEIGTEGFSYIESILICMNE"  
919  
old\_sequence  
/note="g was a in [1]"  
/citation={2}  
922  
old\_sequence  
/note="g was a in [1]"  
/citation={2}

ORIGIN

Alignment Scores:

Pred. No.:	2,578-70	Length:	3543
Score:	1061.50	Matches:	271
Percent Similarity:	50.4%	Conservative:	118
Best Local Similarity:	35.1%	Mismatches:	278
Query Match:	27.4%	Indels:	105
DB:	1	Gaps:	24

US-10-783-417-2 (1-735) x BTTOXD1 (1-3543)

QY 1 MetAaGlnAaSerAaPheAaGluAaIleAaPheSerHisThrSerProTyr 20  
DB 1 ATGAATCCCTTAACAAATTAATAATGATAATGATAATGCTTCACAAAAAAATTA 60  
QY 21 PheProAaArgAaSerAaPheAaPheAaTyrProTyrThrAaAaAaProAaGlnPro 40  
DB 61 -----AATATATCTAAATATATATACAGATATCAATAGAAATAGTCCAAACAAATTA 114  
QY 41 LeuGlnAaThrAaSerAaTyrLeuAaGluAaMetCysGlnGlyAaSerThrGlnTyrGly 60  
DB 115 TTACAAAGGACAAATTAATAAGATGGCTCATATGTGTCAACAGAAATCAGAGATATG 174  
QY 61 AaPheAaPheGluThrPheAaAaSerAaAaPheThrIleAaAaValAaSerAaGlyThrIle 80  
DB ::::::::::::::::::::

DB 175 GGAGATTTTGAACCTTATTGATAGT-----GGTGAACCTGAGCTCATATCATATT 225  
QY 81 ValSerGlyThrLeuLeuAaGlyIleGlyGlyLeuThrSerIleSerGlyProIleGly 100  
DB GTTAGTGGGACCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 273  
QY 101 IleIleGlyAaIleIleIleSerPheGlyThrLeuIleThrValPheTyrProAaGly 120  
DB 274 CTT-----GCTTATATAGTTTGTGATCAATTAATACAGGCTCTTTTCCAGCCCA 324  
QY 121 GluGlnAaPheSerThrValThrThrGlnPheIleIleSerMetGlyGluIlePheValAaPheThr 140  
DB 325 GACCAATCTTAACACA---TGGATGATCTTATTAACACAACTTAATAAATATTAATAAATA 381  
QY 141 ProLeuThrGluSerIleLeuGlnLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 160  
DB 382 GAATATGATCAACAT 441  
QY 161 LeuGlnSerTyrAaThrAaIleLeuAaPheAaPheTyrAaGlyLeuLeuLeuLeuLeuLeu 180  
DB 442 ATCAGCACTTATCATATATCATCTTAATAAATATGAGAG-----AATATATCA 486  
QY 181 GlyLeuProProSerSerAaIleLeuGlnGlnAaIleAaLeuThrLeuLeuLeuLeuLeuLeu 200  
DB 487 AACCCAAATATCTCAGATGTAAGACACAAATCCAGCTAGTTCATTACCATTTTCA 546  
QY 201 AaAaValHisAaAaPheIleAaGluIlePro-----GlyPheGlnLeuGluThrTyr 218  
DB 547 AATGTCAATCCAGAGCTTGTAACTCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 606  
QY 219 LysThrLeuLeuLeuProIleTyrAaGlnAaAaAaPheHisLeuAaLeuGln 238  
DB 607 AACATCTAGAT 666  
QY 239 GlnGlyAaGluLeuAaAaPheGluThrAaAaAaPheHisLeuAaLeuAaPheGluPro 258  
DB 667 CAAGCCCTCAATTTGAAGCTATTTAAATAAACAATCAATTCAGTATTTATTAAGAGCT 726  
QY 259 AaAaAaGlyThrSerAaPheTyrTyrLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 278  
DB 727 ---TTGCCAACAGCAATGAT 783  
QY 279 AaTyrCysAaAaThrTyrAaGlyThrGlyLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 296  
DB 784 AATATATGTATACAACTTATTAATAAATAAAGATTAATTAATAAATAAAGATGATAGT 843  
QY 297 -----AaMetLeuTyrSerIlePheAaAaPheTyrAaGlyTyrMetThrIle 312  
DB 844 AATCTGATGAAAT 903  
QY 313 ThrValLeuAaPheThrIleSerGlnPheSerLeuTyrAaPheIleLeuAaGlyAaPheSer 332  
DB 904 GCTGAT 954  
QY 333 IleGlyGlyIleGluValLeuGlyIleLeuAaGluLeuLeuLeuLeuLeuLeuLeuLeuLeu 352  
DB 955 -----CCATAGGTGTCCAACTGATCAATCTACTCGAAGAAATTAATACAGGTA 999  
QY 353 GluIleAaPheAaPheAaGluLeuProGlnLeuAaGlyAaGlnProAaAaLeuAaThrMetGlu 372  
DB 1000 ---CTTAATCTTGAAGAAAGCCCTTATTAATATATAT-----GACTTCAATATCAAGAG 1050  
QY 373 TyrAaLeuThrAaGlyAaSerPheLeuLeuPheLeuGlnGlnPheIlePheTyr 392  
DB 1051 GATTCATCTTACAGTACGCG---CATTTATTTCTAGGCTTGAATCTTGAATTTTAT 1107  
QY 393 -----ThrGluAaThrAaPheGlyAaAaGly----- 401  
DB 1108 GAAAGAGGCAACATCTCTTAAT 1167  
QY 402 ---LeuValGlyIleSerAaAaPheAaPheAaPheAaPheAaPheAaPheAaPheAaPhe 420  
DB 1168 ACATTTGAT 1227



QY 421 LeuTyrGlyGluArgThrGlySerProThrThrLysThrIleArgProPheGluSerTyr 440  
1228 TTR-----AAATCTGTGTTGGCAACAAATATTATTT----- 1263  
QY 441 LysValSerIleValThrAspArgInsProProValSerProIleGlnProHisPhe 460  
1264 -----TTT 1266  
QY 461 IleIleAsnGlnIleGluLeu-----TyrLeuAsnGlySerSerAsnAsnThr--- 476  
1267 TTAATTAATGTCATAGCTTAGATAATAATATCATTAATATTAATATTACTATAA 1326  
QY 477 -----LeuLysTyrSerIleGly 483  
1327 ATGCAATTTTATTAACCTAAGTACTAGACTTTGGAGAAAGAACTTACAGCAGATCT 1386  
QY 484 SerLeuSerAsnTyrGln---AsnThrThrPhePheGlnPheProArgLysAspCys 502  
1387 GGGCAATTAACCTATAGATGAATATAAATATTGGGTACCAATCTTAAGAGAGA 1446  
QY 503 AsnLeuValIleAspProGlyCysSerProAsnPheAsnAsnTyrSerHisIleLeuSer 522  
1447 GAGAAATCAAGAAACCTACCTTTTCCAAACATATGATTAATCATATTTATATCA 1506  
QY 523 HisPheSerLeuPheThrTyrSerTyrValIleGlyLeuGlnLeuGlnIleLeuAspThr 542  
1507 TTTATTAATAAGTCTTATCCTCGACACATATAAATCTCAAGTAT----- 1554  
QY 543 GlyValIleGlyTyrThrHisSerSerValAspArgTyrAsnAlaIleSerAspLysIle 562  
1555 ---ACGTTTGCTTGACACACTAGCTTGAATCCATAAATATCATTTATACACATTTA 1611  
QY 563 IleThrMecIleProAlaIleLysGlyAsnAsnLeuAspThrAsnSerLysValIleGlu 582  
1612 ACTACCAAAATTCACCTGTAAAGCAATTCACCTGGGCTGCTTAAAGTTGTCTCA 1671  
QY 583 GlyProGlyHisThrGlyGlyAsnLeuValTyrLeuGlnSerGlnIleArgLeuGluIle 602  
1672 GAGCGGTGCATACAGAGGGGATTTAAT---GATTTCAAGATCATTTCAAAATTT 1725  
QY 603 ThrCysGluThrProAsnSerThrGlnSerTyrPheIleArgLeuArgTyrAlaThrAsn 622  
1726 ACATGTCAACACTCAATTTTCAACATGATTTTATAGATTTGTTATGCTTTCAAT 1785  
QY 623 GlyAlaGlyAsnThrLeuProAsnIleSerLeuThrIleProGlyValIleGlyIlePro 642  
1786 GGAAGCGCAAAATCTGAGCTGTATTAATCTTAGATATCCAGGGGAGACAGAACTG--- 1842  
QY 643 ProGlnArgLeuAsnAsnThrPheSerGlyThrAsnTyrAsnAsnLeuGlnTyrGlyAsp 662  
1843 GGTATGCACTCAACCCACTTTTCTGTGACAGATTAATCGAATTTAAATTAAGAAT 1902  
QY 663 PheGlyTyrPheGlnPheProSerThrValThrLeuProLeuAsnArgAsnIleProPhe 682  
1903 TTTCAGTACTTAGAATTTTCTTACAGAGTCAAAATTTGCTCCAAATCAAAATATCTCTT 1962  
QY 683 IlePheAsnArgIleAspVal---SerAsnSerIleLeuIleIleAspLysIleGluPhe 701  
1963 GTGTTTAATCGTTGCGATGTATATACAAACACAGACTACTTATGATTAATTAATTTGAATTT 2022  
QY 702 IleProIleThrSerSerMetHisGlnAsnArgGluLysGlnLysLeuGlnThrIleGln 721  
2023 CTGCCAATTAATCGTTCTATAGAGAGATAGAGAAACAAATAATTAAGAAACAGATCA 2082  
QY 722 ThrLysIleAsnThrPhePheThrAsnHisThrLys 733  
2083 CAATTAATTAATCATTTTATGCAAAATCTATATAA 2118  
Db  
RESULT 5  
LOCUS 108083  
DEFINITION Sequence 1 from Patent EP 0296870.

ACCESSION 108083  
VERSION 108083.1 GI:589204  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 4451)  
AUTHORS Ellar,D.J. and Ward,B.S.  
TITLE New toxin-encoding DNA fragments from *Bacillus thuringiensis*  
JOURNAL subsp. israelensis  
Patent: EP 0296870-A1 1-28-DEC-1998;  
FEATURES location/Qualifiers  
source 1..4451  
/organism="unknown"  
/mol\_type="unassigned DNA"  
ORIGIN  
Alignment Scores:  
Pred. No.: 3,42e-70 Length: 4451  
Score: 1061.50 Matches: 271  
Percent Similarity: 50.4% Conservative: 118  
Best Local Similarity: 35.1% Mismatches: 278  
Query Match: 27.4% Indels: 105  
DB: Gaps: 24  
US-10-783-417-2 (1-735) x 108083 (1-4451)  
QY 1 MetAsnGlnAsnAsnAspAsnAsnGluTyrGluIleIleAspSerHisThrSerProTyr 20  
891 ATGAAATCCCTTATCAAAATATAATGAAATATGAAACATTTAAATGCTTCACAAAAAAATTA 950  
QY 21 PheProAsnArgAsnSerAsnAspSerArgTyrProTyrThrAsnAspProAsnGlnPro 40  
951 -----AATATATCTTAATATATATATACAGATATCCAAATGAAATATGTCACAAACAAATTA 1004  
QY 41 LeuGlnAsnThrAsnTyrLysGluTyrPheAsnMetCysGlnGlyAsnThrGlnTyrGly 60  
1005 TTACAAAGTACAAATATTAATTAAGATGGCTCATATGTGTCAACAGATCAGCAGTATGCT 1064  
QY 61 AspAsnPheGluThrPheAlaSerAlaAspThrIleAlaIleValSerAlaGlyThrIle 80  
1065 GGAAGTTTGAACCTTTATTTATGATAGT-----GGTGAACCTCAGTCCCTATACTATT 1115  
QY 81 ValSerGlyThrLeuLeuAlaGlyIleGlyGlyLeuThrSerIleSerGlyProIleGly 100  
1116 GATGTTGGAGCGCTACTGACTGCTGTTGGGTTCAACA-----CCCTTAGGA 1163  
QY 101 IleIleGlyAlaIleIleIleSerPheGlyThrLeuIleThrValPheThrProAlaGly 120  
1164 CTT-----GCTTTAATAGCTTTTGGTACATTAATACAGTTCTTTTCCAGCCCA 1214  
QY 121 GlnGlnAspLysThrValTyrThrGlnPheIleLysMetGlyGluIlePheValAspThr 140  
1215 GACCAATCTTAACACA---TGAGTGACCTTATTAACACAACTAAATAATTAATAAAAAA 1271  
QY 141 ProLeuThrGlnSerIleLysGlnLeuLysLeuGlnThrLeuGlnGlyPheArgGlnIle 160  
1272 GAAATAGCATCAACATATTAATTAAGTAAATTAATTTTAAACAGTGGTTAATGTT 1331  
QY 161 LeuGlnSerTyrAsnThrAlaLeuAspAspTyrAspGlyLeuLysArgLeuGlnAlaPro 180  
1332 ATCAGCACTTATCATATATACCTTTAAACATGGAG-----AATATCCA 1376  
QY 181 GlyLeuProProSerSerAlaLeuGlnGlnAlaIleAlaLeuThrLeuLysIleArgPheGlu 200  
1377 AACCCAAATAATATCAGAGATGTAAGACAAATCCAGCTAGTTCATTACCATTTTCAA 1436  
QY 201 AsnValHisAsnAspPheIleArgGluIlePro-----GlyPheGlnLeuGlnThrTyr 218  
1437 AATGTCATTCAGAGCTGTAAATCTTGTCTCTTAATCTTAGTGAATTCATTAATTAAT 1496  
QY 219 LysThrLeuLeuLeuProIleTyrAlaGlnAlaIleAsnPheIleLeuAsnLeuGln 238

Db 1497 AACATCACTAGTATTATCTAGTTATGACACAGCAAGCAAACTTACATCTGATCTATTAAT 1556  
 Qy 239 GInGInlAgLuLeuAlaApGlUtrPheAnaIaApIleIseProSeGInIleGluPro 258  
 Db 1557 CAAGCGCGTCAAAATTGAAAGCGATTTAAAAACAATCGACAAATCGATTATTTAAGGCT 1616  
 Qy 259 AenAlaGlyThrSerApeApyrTyrylySLeuLeuLySGLuAnIleProLySrySer 278  
 Db 1617 ---TTGCCAACAGCAATGATTAATATCCAGATTAAGCACTAATAGCATATGAAATTAAC 1673  
 Qy 279 AenTyryCyAlaAenThrTyraTrgThrylyLeuLySAnIleuAryAAspGluPro----- 296  
 Db 1674 AATTAATGTGTACACACTTAATAAAAAAGATTAATTAATTAATAAACAGCGCTGATAGT 1733  
 Qy 297 -----AsnMetLySryPseSerIlePheAenApyrTyraTrgTyrmethThrIle 312  
 Db 1734 AATCTTGATGAAATATATAAAGTGAACACATACAAATAGCTATCGAACAATAATGACTACT 1793  
 Qy 313 ThrValLeuApyrThrIleSeGInPheSerLeuTyraApIleLyAryGlyTyraAAspSer 332  
 Db 1794 GCGATATGATGATGTTGTCACCTCTTCTTAATTAATGATGATGATTAAT----- 1844  
 Qy 333 IleglyGlyIleGluValLyGlyIleLySAnGluLeuThrArgGluIleTyrrThr 352  
 Db 1845 -----CCAAATAGGTGTCCAATCTGAACCTTACTGAGAAATTTATCAGTA 1889  
 Qy 353 GInIleAenPheApyrLeuProGInLeuAryValGInProAenIleuAlaThrmethGlu 372  
 Db 1890 ---CTTAATCTGAGAAAGCCCTTAATAATTAAT-----GACTTCAATATCAAGAG 1940  
 Qy 373 TyraenLeuThrArgIleAseRhePheLySLeuPheSerPheLeuGInPheIlePheTy 392  
 Db 1941 GATTCACCTTACAGTACGACCG--CATTAATTAATCTGATGATCTTGAATTTTAT 1997  
 Qy 393 -----ThrGluAenThrAenPheGlyAenArg----- 401  
 Db 1998 GAAAAAGCCCAACTACTCTTAATAATTTTTCACACCGCACTTAATAATATGTTTCATAC 2057  
 Qy 402 ---LeuValGlyIleSeRAnAryAAspAlaProThrTySerAenThrIleThrGluThr 420  
 Db 2058 ACACCTTGATTAATATATCCAAAAATCTAGTGTGGAATACACATGTAATGATTAATA 2117  
 Qy 421 LeuTyryGlyuAryThrylySeProThrThrylySthrIleAryProPheGluSerTy 440  
 Db 2118 TTA-----AAATCTCTGTTGCGCAAAATATTAAT----- 2153  
 Qy 441 LysValSerIleValThrApyrArgInSerProProValSerProIleGInProHisPhe 460  
 Db 2154 ----- 2156  
 Qy 461 IleIleAenGInIleGluLeu-----TyryLeuAenGlySerSerAenAenThr--- 476  
 Db 2157 TTATTATAATGTCTAATAGCTTAATATAATATCTTAATGATTAATAATTAATTAATA 2216  
 Qy 477 -----LeuLySrySerAlaGly 483  
 Db 2217 ATGATTTTTTTATACTAATAGTACTAGACTTTTGGAGAAAGAACTTACACAGAGATCT 2276  
 Qy 484 SerLeuSerAenTyryGln---AenThrThrPhePheGInPheProAryLySAspCyS 502  
 Db 2277 GGGCAAAATTAATCTTAATGATTAATAATTAATTTTGGGTTACCAATTTCTTAACAGACA 2336  
 Qy 503 AenLeuValIleAAspProGlyCySerProAenPheAenAenTyrySerHisIleLeuSer 522  
 Db 2337 GAGATCAAGAAACCTTACCTTTTCCACATATGATTAATCTTAATGATTAATTTTATTA 2396  
 Qy 523 HisPheSerLeuPheThryrySerTyryValIleGlyLeuGInLeuGInIleLeuApyrThr 542  
 Db 2397 TTTATTAATAAGTCTAGTATCCCTGCAACATATTAATAACAGGTGAT----- 2444  
 Qy 543 GlyValLeuGlyTyrrThrHisSerSerValAAspAryGlyTyraAlaIleSerApyrSile 562  
 Db 2445 ---ACGTTTGCTTGACACACTCTAGTGTGATCTTAATAATCAATTAATTAACACATTTA 2501

Qy 563 IleThrmethIleProAlaIleLyGlyAenAenLeuApyrThrAenSerLySValIleGlu 582  
 Db 2502 ACTACCCCAAAATTCACAGCTGTAAAGCGAATTCATCTGGAGACTGCTTCAAGGTGTGCA 2561  
 Qy 583 GlyProGlyHisThrGlyGlyAenLeuValTyryLeuGInSerGInGlyAryGluGluIle 602  
 Db 2562 GACCTGTGTCTATACAGAGGGGATTTAAT-----GATTTCAAAAGATCATTTCCAAAT 2615  
 Qy 603 ThrCyGluThrProAenSerThrGInSerTyrrPheIleAryGLeuAryTyraIleThraen 622  
 Db 2616 ACATGTCAACACTCAAAATTTTCCAAATCTGATTTTAATAAGATTCGTTATGCTTCAAT 2675  
 Qy 623 GlyAlaGlyAenThrLeuProAenIleSerLeuThrIleProGlyValIleGlyIlePro 642  
 Db 2676 GAAAGCGCAAAATTAATCTGAGCTGTATTAATCTTATGATATCCAGGGGTGACAGAACTG--- 2732  
 Qy 643 ProGInAryGLeuAenAenThrPheSerGlyThryAenTyraAenLeuGInTyryGlyAAsp 662  
 Db 2733 GGTATGGCACTCAACCCCACTTTTCTGTACAGATTAATGCAAAATTAATAAGAT 2792  
 Qy 663 PheGlyTyrrPheGInPheProSerThryValThrLeuProLeuAenAryAAsnIleProPhe 682  
 Db 2793 TTTCACTTCTTAAGAAATTTTCTTAACAGAGTGAATTTGCTCCAAATCAAAACATATCTCT 2852  
 Qy 683 IlePheAenAryAlaAAspVal---SerAenSerIleLeuIleIleApyrLyIleGluPhe 701  
 Db 2853 GGTATTAATCGTTCGATGATATATACAAACACACAGTACTTATGATTAATTAATTAAT 2912  
 Qy 702 IleProIleThrSerSerMetHisGInAenAryGlyGlyGInLySLeuGluThrIleGln 721  
 Db 2913 CTGCCAAATTAATCTGTTCTATTAAGAGATTAAGAGAAACAAAAATTAAGAAACAGTACA 2972  
 Qy 722 ThrLySileAenThrPhePheThraenHisThrLyS 733  
 Db 2973 CAATTAATTAATACATTTATGCAAAATCTTAATAAA 3008  
 RESULT 6  
 LOCUS 109103 4451 bp DNA linear PAT 02-DEC-1994  
 DEFINITION Sequence 1 from Patent WO 8810305.  
 ACCESSION 109103  
 VERSION 109103.1 GI:588188  
 KEYWORDS  
 SOURCE  
 ORGANISM  
 . Unknown.  
 REFERENCE  
 1 (bases 1 to 4451)  
 AUTHORS Ellar,D.J. and Ward,B.S.  
 JOURNAL Patent: WO 8810305-A 1 29-DEC-1988;  
 FEATURES  
 source  
 1..4451  
 /organism="unknown"  
 /mol\_type="unasigned DNA"  
 ORIGIN  
 Alignment Scores:  
 Pred. No.: 3,42e-70 Length: 4451  
 Score: 1061.50 Matches: 271  
 Percent Similarity: 50.4% Conservative: 118  
 Best Local Similarity: 35.1% Mismatches: 278  
 Query Match: 27.4% Indels: 105  
 DB: 6 Gaps: 24  
 US-10-783-417-2 (1-735) x 109103 (1-4451)  
 Qy 1 MetAenGInAenAenAAspAenAenGluTyryGluIleIleAAspSerHisThrySerProTy 20  
 Db 891 ATGAATCCCTTAATCAAAATTAATAATGATATGAAACATTAATGCTTCACAAAAAATTA 950  
 Qy 21 PheProAenAryAAspSerAenAAspSerAryTyrrTyrrThryAenAenProAenGInPro 40  
 Db 951 -----AATTAATATTAATTAATTAATCAAGATATCAATTAAGAAATTAATTAATTA 1004

QY 41 LeuGlnAsnThrAsnTyrIysGluTrpLeuAsnMetCysGlnGlyAsnThrGlnTyrGly 60  
 DB 1005 TTACAAAGACAAATTAATAAGATTGGCTCAATATGTGTCAACAGAAATCAGCAGTATGT 1064  
 QY 61 AspaenpHeGluThrPheAlaSerAlaAspThrIleAlaIleValSerAlaGlyThrIle 80  
 DB 1065 GGAGATTTTGAACCTTTATGTAGT-----GGTGAACCTAGTCTATATCTATT 1115  
 QY 81 ValSerGlyThrLeuLeuAlaGlyIleGlyGlyLeuThrSerIleSerGlyProIleGly 100  
 DB 1116 GTAGTTGGACCGTACTGACTGGTTCCGGTTCCACACA-----CCCTTAGGA 1163  
 QY 101 IleIleGlyAlaIleIleIleSerPheGlyThrLeuIleThrValPheThrProAlaGly 120  
 DB 1164 CTT-----GCTTAATAGCTTTGGTACCTTAATACAGCTCTTTTCAGCCCA 1214  
 QY 121 GlnGlnAspPheThrValIrrpThrGlnPheIleIleIleIleIlePheValAspThr 140  
 DB 1215 GACCAATCAACACA---TGAAGTACTTTATTAACACAATAAATAATTATTAATAA 1271  
 QY 141 ProLeuThrGluSerIleIleGlnLeuIleIleGlnThrLeuGlnGlyPheArgGlnIle 160  
 DB 1272 GAAATAGCATCAACATATTAAGTAAATGCTAAATTAATAATTTTAAACAGTCTTAATGT 1331  
 QY 161 LeuGlnSerTyrAsnThrAlaLeuAspAspTrpArgIleLeuIleArgGlnIlePro 180  
 DB 1332 ATCAGCATTAATCATATATCACTTAACACATGGAG-----AATATCA 1376  
 QY 181 GlyLeuProSerSerAlaLeuGlnGlnAlaAlaLeuThrLeuIleIleArgPheGlu 200  
 DB 1377 AACCCAAATAATCAATCAGATGTAAAGACACAATAATCCAGTACTTATCAATTTCA 1436  
 QY 201 AsnValHisAsnAspPheIleArgGluIlePro-----GlyPheGlnLeuGluThrTyr 218  
 DB 1437 AATGTCATTCAGAGCTGTAAACTTGTCTCTTAATCTAATGATTCAGATTAATCAT 1496  
 QY 219 LysThrLeuLeuLeuProIleTyrAlaGlnAlaAlaAsnPheHisLeuAsnLeuGln 238  
 DB 1497 AACATACAGATATCTAGTAAATGACACAGCAAGCAACTTACATCTGATTAATAAT 1556  
 QY 239 GlnGlyAlaGluLeuAlaAspGluTrpAsnAlaAspIleHisProSerGlnIleGluPro 258  
 DB 1557 CAAGCGTCAAAATTTGACCGATTTTAAATAAACAATCGCAATTCGATTAATTAAGCT 1616  
 QY 259 AsnAlaGlyThrSerAspAspTyrTyrIleIleLeuLeuIleGluAsnIleProLysTyrSer 278  
 DB 1617 ---TTGCCAAGCAGCATTTGATTAATTCAGATTCGATAAGCTATAGACATTAACACT 1673  
 QY 279 AsnTyrCysAlaAsnThrTyrArgThrGlyLeuIleAsnLeuArgAspGluPro----- 296  
 DB 1674 AATTAATGTGTAAACAATTAATAAAGATTAAATTTAATTAAATTAACGACGGCTGATGT 1733  
 QY 297 -----AsnMetLysTrpSerIlePheAsnAspTyrArgArgTyrMetThrIle 312  
 DB 1734 AATCTTGATGAAATTAATTAATGACACATACATACATACATGACAAATAATGACTACT 1793  
 QY 313 ThrValLeuAspThrIleSerGlnPheSerLeuTyrAspIleLysArgTyrArgAspSer 332  
 DB 1794 GCTGATTAATGATGTTGTCACACTTTCTTAATTAATGATGTAGGTAAATAT----- 1844  
 QY 333 IleGlyGlyIleGluValIleGlyIleIleIleIleIleIleIleIleIleIleIleIleIle 352  
 DB 1845 -----CCAAATAGTGTCCAAATCTGAATCTGAACTTCGAAATTAATTAACGCTA 1889  
 QY 353 GluIleAsnPheAspArgLeuProGlnLeuArgValGlnProAsnLeuAlaThrMetGlu 372  
 DB 1890 ---CTTAATCTGAAAGAAAGCCCTTAATAATATAT-----GACTTCAATATCAAGAG 1940  
 QY 373 TyrAsnLeuThrArgAlaSerPheIleLeuPheSerPheLeuGlnGlnPheIlePheTyr 392  
 DB 1941 GATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1997

QY 393 -----ThrGluAsnThrAsnPheGlyAsnArg----- 401  
 DB 1998 GAAAGAGGCAAACTACTCTTAATTAATTTTTCACACAGCACTTAATATATGTTCAATAC 2057  
 QY 402 ---LeuValGlyIleSerAsnAspAlaProThrTyrSerAsnThrIleThrGluThr 420  
 DB 2058 ACACCTGGAATATATATCCCAAAATCTAGTGTGGAAATACAAATGATTAATTAATTA 2117  
 QY 421 LeuTyrGlyIleGlyIleSerProThrThrIleIleArgProPheGluSerTyr 440  
 DB 2118 TTA-----AATCTCTGTGGTGGCAACAATTAATTAAT----- 2153  
 QY 441 LysValSerIleValThrAspArgIleSerProValSerProIleGlnProHisPhe 460  
 DB 2154 -----TTT 2156  
 QY 461 IleIleAsnGlnIleGluLeu-----TyrLeuAsnGlySerSerAsnAsnThr--- 476  
 DB 2157 TTATTAATATGATTAATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2216  
 QY 477 -----LeuLysTyrSerAlaGly 483  
 DB 2217 ATGATTTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2276  
 QY 484 SerLeuSerAsnTyrGln---AsnThrThrPhePheGlnPheProArgLysIleAspCys 502  
 DB 2277 GGGCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2336  
 QY 503 AsnLeuValIleAspProGlyCysSerProAsnPheAsnAsnTyrSerHisIleLeuSer 522  
 DB 2337 GAGATCAAGGAACCTTACCTTTTCCACATATGATTAATTAATTAATTAATTAATTAATTA 2396  
 QY 523 HisPheSerLeuPheThrTyrSerTyrValIleGlyLeuGlnLeuGlnIleLeuAspThr 542  
 DB 2397 TTTATTAATAAGCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2444  
 QY 543 GlyValLeuGlyTyrThrHisSerSerValAspArgTyrAsnAlaIleSerAspLysIle 562  
 DB 2445 ---ACGTTGCTTGACACACTTGTGTTGATCTTAATAATTAATTAATTAATTAATTAATTA 2501  
 QY 563 IleThrMetIleProAlaIleIleGlyIleAsnAsnLeuAspThrAsnSerLysValIleGlu 582  
 DB 2502 ACTACCAAAATTCACGCTGTAAGCAATCACTTGGAGACTGCTTAATTAATTAATTAATTAAT 2561  
 QY 583 GlyProGlyHisThrGlyIleAsnLeuValTyrIleGlnIleGlnGlyArgLeuGluIle 602  
 DB 2562 GGAACCTGCTCATACAGAGGAGATTAAT-----GATTCAAAGATCATTTCAAAAT 2615  
 QY 603 ThrCysGluThrProAsnSerThrGlnSerTyrPheIleArgLeuArgTyrAlaThrAsn 622  
 DB 2616 ACATGTCACACTCAAAATTTTCAACAAATCGTATTTATTAAGATTCGTATGCTCAAAAT 2675  
 QY 623 GlyAlaGlyAsnThrLeuProAsnIleSerLeuThrIleProGlyValIleGlyIlePro 642  
 DB 2676 GGAACCGCAATTAATCTCGAGCTGTTAATATCTTAATTAATTAATTAATTAATTAATTAAT 2732  
 QY 643 ProGlnArgLeuAsnAsnThrPheSerGlyThrAsnTyrAsnAsnLeuGlnTyrGlyAsp 662  
 DB 2733 GGTATGACATCAACCCACTTTTCTGTACAGATTTTACCAATTAATTAATTAATTAATTAAT 2792  
 QY 663 PheGlyTyrPheGlnPheProSerThrValThrLeuProLeuAsnArgAsnIleProPhe 682  
 DB 2793 TTTCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2852  
 QY 683 IlePheAsnArgAlaAspVal---SerAsnSerIleLeuIleIleAspLysIleGluPhe 701  
 DB 2853 GTGTTTAATCGTGTGATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2912  
 QY 702 IleProIleThrSerSerMetHisGlnAsnArgIleIleGlyIleGlyIleGlyIleGln 721  
 DB 2913 CTGCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2972  
 QY 722 ThrLysIleAsnThrPhePheThrAsnHisThrLys 733



```

Db      1347 -----CCAATAGGTGTCATCTGAATCTGAACTTACTGAGAAATTTATCAGGTA 1391
Qy      353 GIUIleAenPheAspArgLeuProGlnLeuArgValGlnProAsnLeuAlaThrMetGlu 372
Db      1392 ---CTTAACCTCGAAGAAAGCCCTATTAATAATATAT- ---GACTTCAATATCAAGAG 1442
Qy      373 TyrAsnLeuThrArgAlaSerPheGlySerPheGlyLeuGlnPheIlePheTyr 392
Db      1443 GATTCACCTTACACGTAGACGCG---CATTTATTTACTGCGCTGATCTTGAATTTTAT 1499
Qy      393 -----ThrGluAsnThrAsnPheGlyAsnArg----- 401
Db      1500 GAAAAAGCGCAAACTACTCTCTAATATTTTTCACAGCATTAATATATGTTTCATATC 1559
Qy      402 ---LeuValGlyIleSerAsnArgAspAlaProThrTyrSerAsnThrIleThrGluThr 420
Db      1560 ACACCTGATTAATATATATCCCAAAATCTAGTGTGTTTGGAAATCACAATATCACTGATTA 1619
Qy      421 LeuTyrGlyIuArgThrGlySerProThrThrLysThrIleArgProPheGluSerTyr 440
Db      1620 TTA-----AAATCTTGTTGGTGGACAAATATTTATAT- 1655
Qy      441 LysValSerIleValThrAspArgInsProProValSerProIleGlnProHisPhe 460
Db      1656 -----TTT 1658
Qy      461 IleIleAsnGlnIleGluLeu-----TyrLeuAsnGlySerSerAsnThr--- 476
Db      1659 TTATTAATATGTCATAGCTTAGATTAATAATATCTAAATGATTAATATATTTACTTAA 1718
Qy      477 -----LeuLysTyrSerIleGly 483
Db      1719 ATGGAATTTTATAAATAAGTACTAGACTTTTGGAGAAAGACTTACAGAGATCT 1778
Qy      484 SerLeuSerAsnTyrGln---AsnThrThrPhePheGlnPheProArgLysIleAspCys 502
Db      1779 GGGCAAAATACCTATGATGATTAATAAATATTTTGGGTACCAATCTTAAACCAAGA 1838
Qy      503 AsnLeuValIleAspProGlyCysSerProAsnPheAsnAsnTyrSerHisIleLeuSer 522
Db      1839 GAGAAATCAAGAAACCTACCTTTTCCAAACATGATGATGATGATGATGATGATGATGAT 1898
Qy      523 HisPheSerLeuPheThrTyrSerTyrValIleGlyLeuGlnLeuGlnIleLeuAspThr 542
Db      1899 TTTATTAATAAGTCTTATGATCCCTGCAACATATTAATACTCAAGTGTAT- 1946
Qy      543 GlyValLeuGlyTyrThrHisSerSerValAspArgTyrAsnAlaIleSerAspLysIle 562
Db      1947 ---ACGTTGCTTGGACACACTGATGATGATGATGATGATGATGATGATGATGATGATGAT 2003
Qy      563 IleThrMetIleProAlaIleLysGlyAsnAsnLeuAspThrAsnSerLysValIleGlu 582
Db      2004 ACTACCCCAATTCACAGCTGTAAGCGAATTCACCTGGAGACGCTTCTAAGGTGTTCAA 2063
Qy      583 GlyProGlyHisThrGlyLysLeuValTyrLeuGlnSerGlnLysArgLeuGluIle 602
Db      2064 GGAACCTGTCATACAGAGAGGAGATTAAAT- ---GATTTCAAAATCATTTCAAAAT 2117
Qy      603 ThrCysGluThrProAsnSerThrGlnSerTyrPheIleArgLeuArgTyrAlaThrAsn 622
Db      2118 ACATGTCACACTCAAAATTTTCAACATGATGATTTTATAGATTTGTTATGCTTCAAT 2177
Qy      623 GlyAlaGlyAsnThrLeuProAsnIleSerLeuThrIleProGlyValIleGlyIlePro 642
Db      2178 GGAAGGCAAAATATCTGAGCTGTATTAATCTTAGATCCAGGGGTAGACAGAACTG- 2234
Qy      643 ProGlnArgLeuAsnAsnThrPheSerGlyThrAsnTyrAsnAsnLeuGlnTyrGlyAsp 662
Db      2235 GGTATGCACTCAACCCACTTTTGTGACGATTAATACGAATTAATAATTAAGAT 2294
Qy      663 PheGlyTyrPheGlnPheProSerThrValThrLeuProLeuAsnArgAsnIleProPhe 682
Db      2295 TTTCAGTACTTAGAATTTTCTTAACGAGGTGAATTTGTCCTCAAAATCAAAACATATCTCT 2354

```

```

Qy      683 IlePheAsnArgAlaAspVal---SerAsnSerIleLeuIleIleAspLysIleGluPhe 701
Db      2355 GGTATTAATCTGTCGAGATGATATACAAACACACAGTACTTATGATTAATTTGAATTT 2414
Qy      702 IleProIleThrSerSerMetHisGlnAsnArgIuLysGlnLysLeuGlnThrIleGln 721
Db      2415 CTGCAATTAATCTCGTTCTATTAAGAGGATAGAGGAAACAAAAATTTAGAAACAGTACAA 2474
Qy      722 ThrLysIleAsnThrPhePheThrAsnHisThrLys 733
Db      2475 CAATTAATTAATATCAATTTATGCAAAATCTATTAATA 2510

RESULT 8
E01676 4934 bp DNA linear PAT 29-SEP-1997
LOCUS E01676
DEFINITION DNA sequence encoding Bacillus thuringiensis insectivorous protein.
ACCESSION E01676.1 GI:2169929
VERSION JP 1988230090-A/1.
KEYWORDS Bacillus thuringiensis
SOURCE Bacillus thuringiensis
ORGANISM Bacillus thuringiensis
Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus; Bacillus cereus group.
REFERENCE 1 (bases 1 to 4934)
AUTHORS Komano, T. and Himeno, M.
TITLE INSECTICIDAL PROTEIN OF STRAIN BACILLUS THURINGIENSIS VAR JOURNAL Patent: JP 1988230090-A 1 26-SEP-1988;
COMMENT SUMITOMO CHEM CO LTD
OS Bacillus thuringiensis
PN JP 1988230090-A/1
PD 26-SEP-1988
PF 19-MAR-1987 JP 1987066844
PI KOMANO TORU, HIMENO MICHIO
PC C12N15/00, A01N63/00, C12N1/20, C12P21/02, A01N63/02, (C12N1/20, C12R1:19);
CC CC
CC topology: Linear;
CC hypothetical: No;
CC anti-sense: No;
CC *source: strain=HD522;
FH Key Location/Qualifiers
FT 5'UTR 621..1368
FT CDS 1369..4911
FT FT /product='Bacillus thuringiensis insectivorous protein'.
FEATURES
source location/Qualifiers
1..4934
/organism='Bacillus thuringiensis'
/mol_type='genomic DNA'
/db_xref='taxon:1428'

ALIGNMENT Scores:
Pred. No.: 4,63e-70 Length: 4934
Score: 1060.50 Matches: 271
Percent Similarity: 50.4% Conservative: 118
Best Local Similarity: 35.1% Mismatches: 278
Query Match: 27.4% Indels: 105
Gaps: 24
DB: 6

US-10-783-417-2 (1-735) x E01676 (1-4934)
Qy      1 McLaNGInAsnAspAsnAsnGluTyrGluIleIleAspSerHisThrSerProTyr 20
Db      1369 ATGAACTCTTATCAAAATTAATAATGATATGAAACATTAAGCTTCACAAAAAAATTA 1428
Qy      21 PheProAsnArgAsnSerAsnAspSerArgTyrProTyrThrAsnAsnProAsnGlnPro 40
Db      1429 -----AATATATCTTAATATTAATATCAAGATATCAAGAAATTAAGTCCAAACATTA 1482
Qy      41 LeuGlnAsnThrAsnTyrLysGluThrPheAsnMetCysGlnGlyAsnThrGlnTyrGly 60

```



```
RESULT 9
BTPROTOXIS/c 127923 bp DNA linear BCT 16-APR-2005
LOCUS Bacillus thuringiensis subsp. israelensis plasmid pBtoxis.
DEFINITION AL731825
VERSION AL731825.1 GI:21685410
KEYWORDS
SOURCE
ORGANISM Bacillus thuringiensis serovar israelensis
Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus; Bacillus
cereus group.
REFERENCE
AUTHORS 1 (bases 1 to 127923)
Berry,C., O'Neill,S., Ben-Doy,E., Jones,A.F., Murphy,L., Quail,M.A.,
Holden,M.T., Harris,D., Zaritsky,A. and Parkhill,J.
TITLE Complete sequence and organization of pBtoxis, the toxin-coding
plasmid of Bacillus thuringiensis subsp. israelensis
JOURNAL Appl. Environ. Microbiol. 68 (10), 5082-5095 (2002)
PUBMED 12343359
REFERENCE 2 (bases 1 to 127923)
Parkhill,J.
TITLE Direct Submission
AUTHORS Submitted (19-APR-2002) Submitted on behalf of the pBtoxis
sequencing team, Sanger Centre, Wellcome Trust Genome Campus,
Hinxton, Cambridge CB10 1SA E-mail: parkhill@sanger.ac.uk
JOURNAL
COMMENT Details of pBtoxis sequencing at the Sanger Centre are available on
the World Wide Web.
FEATURES
source location/Qualifiers
1. 127923
/organism="Bacillus thuringiensis serovar israelensis"
/mol_type="genomic DNA"
/db_xref="taxon:1430"
/note="Plasmid pBtoxis"
1145. 1149
RBS
1163. 2089
gene
1163. 2089
CDS
/gene="pBt001"
/note="pBt001"
/note="Similar in part to Bacillus anthracis pXol-49
TR:Q9X319 (EMBL:AF065404) (227 aa) fasta scores: E():
8.9e-44, 78.48% id in 158 aa"
/codon_start=1
/transl_table=11
/protein_id="CAD30064.1"
/db_xref="GI:21685411"
/db_xref="UniProt/TREMBL:O8KNX5"
/translation="MSTHLNFSQMPKVGQVFPDGLFNIITLSENRKEIPSSALATY
IMLHPCNDIGMLPRRFQIMDLAKSGIPYTTITGFOVCLERKIVREIPGNRTIYE
IVDALVYHNTAAETDDMINISLSYRIPPLYLTUTLLSLVARNRNGIMLLELINT
FSRKIGMEHYKHIKIDFETTRQAPLKEKLNNAKRVQYEIVIPIKENAVDLMDCK
SSESRITRVRQOVQIIEKFNIVSSNCVIENDKKEIHRPIAKCRKSAVRLKMGQ
ALRKDKNINIMAFQGEIYDIIEIYIPTQKQN"
2246. 3106
repeat_unit
/note="IS240"
complement(2246. 2262)
repeat_unit
/rpt_type=INVERTED
2337. 3044
gene
2337. 3044
CDS
/gene="pBt003"
2337. 3044
/gene="pBt003"
/note="pBt003"
/note="Similar to Bacillus thuringiensis insertion element
IS240-a protein TR:Q45766 (EMBL:M23740) (235 aa) fasta
scores: E(): 2.2e-92, 99.57% id in 235 aa, and to
Mycobacterium fortuitum, transposase tnp tnp6100 TR:Q49185
(EMBL:X53635) (254 aa) fasta scores: E(): 1.4e-37, 48.05%
id in 231 aa"
/codon_start=1
/transl_table=11
/product="insertion element IS240 protein"
/protein_id="CAD30065.1"
/db_xref="GI:21685412"

misc_feature
/db_xref="GOA:O8KNX4"
/db_xref="InterPro:IPR001584"
/db_xref="UniProt/TREMBL:O8KNX4"
/translation="MEKENIFKMKHYQADMILWVRYLRYNLSFRDLVEMMERGLS
LSHTITMRWVHQYGPBELNERIRKHLKRTNDSMRVDEYIKIKGENMYLYRAVDSGNT
LDPLYSKRRDAKAAKCFLLKALASPHVLPKPVITVDGKAYPAVALRELKNEKSIYGM
PLRVKYLNNMIEQDHPFKRRILNMLGLKSMQTAVKMIAGIEAMHWKQQLKLRQO
SAONQNRCTIHLFGTLTA"
2538. 2933
misc_feature
/gene="pBt003"
/note="Hmmpfam hit to PF00665, Integrase core domain"
3090. 3106
repeat_unit
/rpt_type=INVERTED
3521. 4381
repeat_unit
/note="IS240"
complement(3521. 3535)
repeat_unit
/rpt_type=INVERTED
3612. 4319
gene
3612. 4319
CDS
/gene="pBt004"
3612. 4319
/note="pBt004"
/note="Similar to Bacillus thuringiensis insertion element
IS240-a protein TR:Q45766 (EMBL:M23740) (235 aa) fasta
scores: E(): 3.5e-91, 99.14% id in 235 aa, and to
Mycobacterium fortuitum, transposase tnpa or tnp6100
TR:Q49185 (EMBL:X53635) (254 aa) fasta scores: E():
1.1e-37, 48.05% id in 231 aa"
/codon_start=1
/transl_table=11
/product="insertion element IS240 protein"
/protein_id="CAD30066.1"
/db_xref="GI:21685413"
/db_xref="GOA:O8KH55"
/db_xref="InterPro:IPR001584"
/db_xref="UniProt/TREMBL:O8KH55"
/translation="MEKENIFKMKHYQADMILWVRYLRYNLSFRDLVEMMERGLS
LSHTITMRWVHQYGPBELNERIRKHLKRTNDSMRVDEYIKIKGENMYLYRAVDSGNT
LDPLYSKRRDAKAAKCFLLKALASPHVLPKPVITVDGKAYPAVALRELKNEKSIYGM
PLRVKYLNNMIEQDHPFKRRILNMLGLKSMQTAVKMIAGIEAMHWKQQLKLRQO
SAONQNRCTIHLFGTLTA"
3813. 4208
misc_feature
/gene="pBt004"
/note="Hmmpfam hit to PF00665, Integrase core domain"
4365. 4381
repeat_unit
/rpt_type=INVERTED
complement(4634. 5275)
gene
/gene="pBt005"
complement(4634. 5275)
CDS
/gene="pBt005"
/note="Similar to N-terminus of Bacillus anthracis pXol-18
TR:Q9X2Y9 (EMBL:AF065404) (315 aa) fasta scores: E():
4.2e-56, 84.15% id in 183 aa, and to Bacillus halodurans
Bh2364 protein TR:Q9KAC5 (EMBL:AP001515) (378 aa) fasta
scores: E(): 1.6e-18, 35.45% id in 189 aa, and weakly to
Lactobacillus delbrueckii integrase/recombinase orf2
TR:Q48538 (EMBL:Z50864) (333 aa) fasta scores: E(): 6.3,
28.88% id in 90 aa, and to Bacillus thuringiensis
recombinase tnpI SW:TNRI BACTV (P10020) (284 aa) fasta
scores: E(): 8.5, 23.88% id in 180 aa"
/codon_start=1
/transl_table=11
/product="Integrase/recombinase family protein"
/protein_id="CAD30067.1"
/db_xref="GI:21685414"
/db_xref="GOA:O8KNX3"
/db_xref="InterPro:IPR004107"
/db_xref="UniProt/TREMBL:O8KNX3"
/translation="MSKNINLERTLAASAVLIRKGRKSGSTIKRYAVYIDQPYKRLA
NKKLNIHKSAGSELSADYQYTFSELEDGRKSQKTRRHIVVYLUKRLHPLCLVSLDG
INLSLIPQSLNDNDFTMEWEKLLKQVLSITKGLTEQAKYRPLIMRANVCILWV
NYGSLQELVSLNNSHINPARNICVYKIKIDTLNRDILFVQLSEIVAFHRSF"
complement(4985. 5251)
/misc_feature
/gene="pBt005"
```

```

/note="TMMPfam hit to PF02899, Phage integrase, N-terminal SAM-like domain"
complement (5283. .5286)
complement (5364. .5507)
/gene="pBC006"
complement (5364. .5507)
/gene="pBC006"
/note="Similar to Bacillus anthracis pxc1-17 TR:O9XZY8 (EMBL:AF065404) (47 aa) fasta scores: E(): 2.1e-12, 68.08% id in 47 aa"
/codon start=1
/transl table=1
/product="putative integral membrane protein"
/protein_id="CAD30068.1"
/db_xref="GI:21685415"
/db_xref="UniProt/TrEMBL:O8KNX2"
/translacion="MAVLKDSNYTEMVYIEGFKGIIFWGLIFVVPFWSIMATFIW LCK"
complement (5370. .5435)
/gene="pBC006"
/note="1 probable transmembrane helix predicted for pBC006 by TMHMM2.0"
complement (5515. .5519)
complement (6451. .8160)
/gene="pBC007"
complement (6451. .8160)
/gene="pBC007"
/note="pBC007"
/note="Similar to Bacillus anthracis pxc1-16 TR:O9XZY7 (EMBL:AF065404) (569 aa) fasta scores: E(): 0.96.13% id in 569 aa, and to Bacillus thuringiensis pxc1 orf16-like protein TR:CA050562 (EMBL:AJ296638) (310 aa) fasta scores: E(): 6.5e-122, 99.67% id in 310 aa"
/codon start=1
/transl table=1
/product="conserved hypothetical protein"
/protein_id="CAD30069.1"
/db_xref="GI:21685416"
/db_xref="UniProt/TrEMBL:O8KNX1"
/translacion="MGKSFIIIONRSLNGKTTAAVVVTLKSWQFVAFAFEKYSYKD TYLDDILQHTVEERLADVEGFVSKAPSPSRAMSTLAEKTRKHLGIYNADLEV FYLHKKCFNGTVIPNVTIHMMEDYKQKSEFPAYIOHSQFYALKKLNHNISLEN GLDRYTIKLPKNEBTEKRNPPVYISPIITPKDFKLSIAKKLFDIMNQHTET TLKSLDKQDBRGKHTFGGNTFRLEKTYPHQIRAVYELTALPCTGNPLFKICMO KGVGTRKRYTTLYLHISDFLCSKEASEBQHPFPKATYAKAFIEAVIQENNG ELSDNNKFIHLVLYKTCROIIRSYIRGLDMVDEKSGYPIKLYTLTKLIMQTSQYI LDTAKEGIVPLIAOHIPKERSNREDAVNPFGIAHSMYSLHNKMFNVAHALKOK PAVYVTEESYRNTLYKQERTLPKXYAYDQGVLAHYALLETIMPEKRLVNHKERTI PSQVRFNFIETLIDLPQEKRLVIELPQFMLEPFRMTFELRYNAGVTTTAPOVTHAM EIK"
complement (8168. .8173)
RBS complement (8168. .8173)
RBS 8947. .8951
RBS 8959. .10653
gene /gene="pBC009"
8959. .10653
RBS /gene="pBC009"
/note="Similar to Bacillus anthracis pxc1-14 TR:O9XZY5 (EMBL:AF065404) (564 aa) fasta scores: E(): 1.6e-191,
Alignment Scores:
Pred. No.: 2.73e-68 Length: 127923
Score: 1060.50 Matches: 271
Percent Similarity: 50.4% Conservative: 118
Best Local Similarity: 35.1% Mismatches: 278
Query Match: 27.4% Indels: 105
DB: 1 Gaps: 24
US-10-783-417-2 (1-735) x BTPBTOXIS (1-127923)
QY 1 MetAENGlnAaenAaenAGLUtYcYulIeIaepSerHisThrserProtyr 20
Db ATGAATCTTATCAAAATATAAATGAATATGAAACATTAAATGCTTCACAAAAAATTA 96465
QY 21 PhepRoAenAgaenSerAenApsSerAArgtyrProtyrThraAnsnProAengInPro 40

```

Db 96468 -----AATATATCTAATATATTTATACAGATATCCATAGAAAATATGTCCAAAACATTA 96415  
OY 41 LeuGlnAsnThrAsnTyrLysGluTrpLeuAsnMetCysGlnGlyAsnThrGlnTyrGly 60  
Db 96414 TTACAAAGTCAATATTAATAAGATGGCTCAATATGTGTCAACAGAAATCAGCAGTATGGT 96355  
OY 61 AspAsnPheGluThrPheAlaSerAlaAspThrIleAlaValSerAlaGlyThrIle 80  
Db 96354 GGAATTTTAAACTTTATATGATAG-----GGTGAACCTCAGTGCCTATACTATT 96304  
OY 81 ValSerGlyThrLeuLeuAlaGlyIleGlyGlyLeuThrSerIleSerGlyProIleGly 100  
Db 96303 GTATTTGGAGCCGACTGACTGCTTCCGAGTTCAACACA-----CCCTTAGGA 96255  
OY 101 IleIleGlyAlaIleIleIleSerPheGlyThrLeuIleThrValPheTrpProAlaGly 120  
Db 96255 CTT-----GCTTTAATAGGTTTGGTACATTAATACAGTTCTTTTTCAGGCCAA 96208  
OY 121 GluGlnAspLysThrValTyrTrpGlnPheIleLysMetGlyGluIlePheValAspThr 140  
Db 96204 GACCAATCTAACA-----TGAAGTGACTTATATACAAACCTAAATAATTTATAAAAA 96144  
OY 141 ProLeuThrGluSerIleLysGlnLeuLysLeuGlnThrLeuGluGlyPheArgGlnIle 160  
Db 96147 GAAATATACATCAACATATATAGTAATGCTAATATAATTTTAAACAGTCTGTTAAATGTT 96088  
OY 161 LeuGlnSerTyrAsnThrAlaLeuAspAspTyrArgLysLeuLysArgLeuGlnAlaPro 180  
Db 96087 ATCAGCACTTATCTATATCACTACCTTAAACATGGAG-----AATATCCA 96044  
OY 181 GlyLeuProSerSerAlaLeuGlnGlnAlaAlaLeuThrLeuLysIleArgPheGlu 200  
Db 96042 AACCCACAAATATCTCAGATGTAGAGACCAATCCAGCTAGTACTTATTCACATTTTCAA 95983  
OY 201 AsnValHisAsnAspPheIleArgGluIlePro-----GlyPheGlnLeuGluThrTyr 218  
Db 95982 AATGTCAATCCAGAGCTTGAACTCTGTCTCTCTTAATCTTAGTAATGGCATTACTAT 95922  
OY 219 LysThrLeuLeuLeuProIleTyrAlaGlnAlaAlaAsnPheIleLeuAsnLeuGln 238  
Db 95922 AACCTACTAGATTATTACTATTAATGCAAGCAGCAAACTTACATGCTAGTAATTAAT 95866  
OY 239 GlnGlyAlaGluLeuAlaAspGluTrpAsnAlaAspIleHisProSerGlnIleGluPro 258  
Db 95862 CAACCCGTCAATTTGAAGCGTATTTAAAAAACAATGCACAAATTCATTATTAGAGCCT 95803  
OY 259 AsnAlaGlyThrSerAspAspTyrTyrLysLeuLeuLysGluAsnIleProLysTyrSer 278  
Db 95802 ---TTGCCAAGCAGCAATGATTTATTCACGTTTGACTTAACCTTACAGAAATTAACCT 95746  
OY 279 AsnTyrCysAlaAsnThrTyrArgThrGlyLeuLysAsnLeuArgAspGluPro----- 296  
Db 95745 AATTATTTGTACAACTTATATAAAAAAGATTAAATTAATTTAAAAACAGCGCTGATAGT 95688  
OY 297 -----AsnMetLysTrpSerIlePheAsnAspTyrArgArgTyrMetThrIle 312  
Db 95685 AATCTTGATGGAAATATAAAGCTGGAACACATACATACGATCGAACAATAAAGTACTACT 95622  
OY 313 ThrValLeuAspThrIleSerGlnPheSerLeuTyrAspIleLysArgTyrArgAspSer 332  
Db 95625 GCTGTATTAGATCTGTGCACTCTTTCCTTAATTAATGATGTAGATAATAT----- 95575  
OY 333 IleIleGlyAlaIleGluValLysGlyIleLysAsnGlnLeuThrArgGluIleTyrThrThr 352  
Db 95574 -----CCAAATGAGTGTCCAATCTGAACCTTACTGCAAAAATTTTACAGGTA 95533  
OY 353 GluIleAsnPheAspArgLeuProGlnLeuArgValGlnProAsnLeuAlaThrMetGlu 372  
Db 95529 ---CTTAACCTTCGAAGAAAGCCCTTAATAATATAT-----GACCTTCAATATCAAGAG 95479  
OY 373 TyrAsnLeuThrArgAlaSerPheLysLeuPheSerPheLeuGlnGlnPheIlePheTyr 392



```

Db 95478 GATTCACCTTACAGTAGACGCG---CAATTATTTACTGGCTGATCTTGGAAATTTTAT 95422
Qy 393 -----ThrgluAenThrAsnphgGlyAsnArg----- 401
Db 95421 GAAAAAGCCAAACTACCTCTTAATATTTTTCACAGCAATTAATATGTTCACTTAC 95362
Qy 402 ---LeuValGlyIleSerAsnArgAspAlaProThrTySerAsnThrIleThrGluThr 420
Db 95361 ACACCTGATATATATATCCAAAAATCTAGGTTTTTGGAAATCACAACTGATCTATAA 95302
Qy 421 LeuTyGlyGluArgThrGlySerProThrLysThrIleArgProphgGluSerTyR 440
Db 95301 TTA-----AAATCTGTGGTTGGCAACAAATATTATATT----- 95266
Qy 441 LysValSerIleValThrAspArgInsSerProValSerProIleGlnProHisPhe 460
Db 95265 -----TTT 95263
Qy 461 IleIleAsnGlnIleGluLeu-----TyLeuAsnGlySerSerAsnAsnThr--- 476
Db 95262 TTATTAAATGCTAGCTTAGCTTAATATATATCTAAATGATTATATATATTAGTAAA 95203
Qy 477 -----LeuLysTySerIleGlyGly 483
Db 95202 ATGATTTTATTAATACTAGTACTAGTACTAGTCTTTGGAGAAAGAACTTACAGCAGGATCT 95143
Qy 484 SerLeuSerAsnTyrcIn---AsnThrThrPhePheGlnPheProArgLysAspCys 502
Db 95142 GGGCAAAATTAATATATATGATGTAATATAAATATTTTGGGTTACCAATCTTAAACAGAGA 95083
Qy 503 AsnLeuValIleAspProGlyCysSerProAsnAsnAsnTySerHisIleLeuSer 522
Db 95082 GAGAAATCAAGAAACCTTACCTTTCCACATATGATATACATATGATATTTATATCA 95023
Qy 523 HisPheSerLeuPheThrTySerTyValIleGlyLeuGlnIleLeuAspThr 542
Db 95022 TTTATTAATAAGCTTAGTATCCCTGCACATATATAAACAAGTGTAT----- 94975
Qy 543 GlyValLeuGlyTyThrHisSerSerValAspArgTyAsnAlaIleSerAspLysIle 562
Db 94974 ---ACGTTGCTGGACACACTTACGTTGATGCTTAAATAATCAATTTATACACATTTTA 94918
Qy 563 IleThrMetIleProAlaIleLysGlyAsnAsnLeuAspThrAsnSerLysValIleGlu 582
Db 94917 ACTACCAAAATTCACCTGTAAGAAAGCAATTCCTGGAGCTGCTTCTAAGGTTGTTCA 94858
Qy 583 GlyProGlyHisThrGlyLysAsnLeuValTyLeuGlnSerGlnGlyArgLeuGluIle 602
Db 94857 GGAACCTGGTCATACAGAGGAGGATTAATT-----GATTTCAAGATCATTTCAAAATT 94804
Qy 603 ThrCysGluThrProAsnSerThrGlnSerTyPheIleArgLeuArgTyAlaThrAsn 622
Db 94803 ACATGTCAACACTCAAAATTTTCAACAATGATGATTTATTAAGAATTCGTTATGCTTCAAAAT 94744
Qy 623 GlyAlaGlyAsnThrLeuProAsnIleSerLeuThrIleProGlyValIleGlyIlePro 642
Db 94743 GGAAGCCAAATATCTGAGCTGTTATTAATCTTAGATCCAGGGGTAGCAGAACTG--- 94687
Qy 643 ProGlnArgLeuAsnAsnThrPheSerGlyThrAsnTyAsnAsnLeuGlnTyArgLysP 662
Db 94686 GGTATGCACTCAACCCCACTTTTCTGTGGTACAGATTATACGAATTTAAATAAAGAT 94627
Qy 663 PheGlyTyPheGlnPheProSerThrValThrLeuProLeuAsnArgAsnIleProPhe 682
Db 94626 TTTCACTACTTGAATTTTCTAAGAGGTGAATTTGCTCCAATCAAAACATATCTCTT 94567
Qy 683 IlePheAsnArgAlaAspVal---SerAsnSerIleLeuIleIleAspLysIleGluPhe 701
Db 94566 GGTATTAAATGCTTGGATGATATATACAAACACAGTACTTATTAATAAATTTGAATTT 94507
Qy 702 IleProIleThrSerSerMetHisGlnAsnArgGlnLysGlnLysLeuGlnThrIleGln 721
Db 94506 CTGCCAATTAATCTGTTCTATTAAGAGAGATAGAGAAACAAATAATTTAGAAACAGTACAA 94447

```

```

Qy 722 ThrLysIleAsnThrPhePheThrAsnHisThrLys 733
Db 94446 CAAATATTAATTAATACATTTTATGCAAAATCTCTATATAA 94411

RESULT 10
BACMSOB
LOCUS 3753 bp DNA linear BCT 26-APR-1993
DEFINITION B.thuringiensis insecticidal endotoxin gene, complete cd.
ACCESSION M12662
VERSION M12662.1 GI:143228
KEYWORDS
SOURCE
ORGANISM
Bacillus thuringiensis
Bacillus thuringiensis
Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus; Bacillus
cereus group.
1 (bases 1 to 3753)
AUTHORS
Thorne, L., Garduno, F., Thompson, T., Decker, D., Zouner, M., Wild, M.,
Walfeld, A.M. and Pollock, T.J.
TITLE
Structural similarity between the lepidoptera- and diptera-specific
insecticidal endotoxin genes of Bacillus thuringiensis subsp.
'kurstaki' and 'israelensis'
JOURNAL
J. Bacteriol. 166 (3), 801-811 (1986)
PUBMED
3011746
COMMENT
source text: B.thuringiensis subspecies israelensis DNA.
FEATURES
location/Qualifiers
source 1..3753
/organism="Bacillus thuringiensis"
/mol_type="genomic DNA"
/db_xref="taxon:1428"
941..2968
/note="insecticidal endotoxin (put.); putative"
/codon_start=1
/cran1_table=11
/protein_id="AAA22614.1"
/translation="MNPKYONKNEYEPNAPSNGFSKSNYSRPLANKRNOPLKNTNY
KDLNVCODNOOYGNAGNAPASERTIVGAGIIVGTMGLAFAPVLAAGISRGTL
LPFQSGDPRANWQDLNIGRPIQEDIKKINIVTSIVPIKQDLKYGQFPKWE
PARTHANKAAVHDLFTTLEPIIDKOLDKNNASYRIPLPAVAGIATWHLKRAA
TYNNTWLNQGINSTFNSNYYQGLKRLKIOBYDYCIQTNAGLTMIRITNATWN
MYNTIRLEMTLTVLDLIAIPNYDEKPIGVSKELIBEVYNNVNSDFERTITLENG
LTRNPTLFWINQGFYRNSRDLDPYDIFSTQNMFAFTINDRNIICGAVHGNL
ISQTSKPLPPYRNKPIQKIVREKSDIIVEMIFPSNSSEVRYSNSRTIENNYK
RTDSYMIKQTKRNEYGHTLSYITDNYIRSVARRRVAASWTHTSDPONTDID
NITQIHAKLVSSDSKIVKPGHTGDLVILKQSDMPKRVFLKNSKQIVRIRYA
TNAPKTVFLTGIDIPISVELPSTSRQNPNAIDLYADFGVYTPRPVKNKTFEGEDT
ILMTLYGTPNHSYNIYIDKIERIPITQSVLDYTEKONIEKTQKIVMDLFPV"

ORIGIN
Alignment Scores:
Pired. No.: 7,08e-58 Length: 3753
Score: 897.50 Matches: 247
Percent Similarity: 48.3% Conservative: 126
Best Local Similarity: 32.0% Mismatches: 270
Query Match: 23.2% Indels: 129
DB: 1 Gaps: 30

US-10-783-417-2 (1-735) x BACMSOB (1-3753)
Qy 1 MetAngGlnAsnAsnAspAsnAsnGluTyArgIleIleAspSerHisThrSerProTyR 20
Db 941 ATGAATCCATATCAAAATTAAGATGATATGAAATATTCAATGCTTCATCAATGTTT 1000
Qy 21 PheProAsnArgAsnSerAsnAspSerArgTyProTyThrAsnAsnProAsnGlnPro 40
Db 1001 -----AGCAAGCTTAATTAATTAATCTAGATATCATTAAGCAATTAAGCAATCAACA 1054
Qy 41 LeuGlnAsnThrAsnTyArgIleGluThrLeuAsnMetCysGlnGlyAsnThrGlnTyArg 60
Db 1055 CTGAAATAACCAAAATTAACAAATATGCTCAATGCTGTCAGATATTAACAAATATATGCTC 1114
Qy 61 AspAsnPhgGluThrPheAlaSerAlaAspThrIleAlaValSerAlaGlyThrIle 80

```

Db 1115 AATAATGGCGGGAATTTCCTAGTCTGTAACATATGTTGGAGTATGAGGTATTTAT 1174  
QY ValSerGlyThrLeuLeuAlaGlyIleGlyGlyLeuThrSerIleSerGlyProIleGly 100  
Db 1175 GTAAGAGAACTATGTTAGA-----GCTTTGGCTGCCCT----- 1210  
QY 101 IleIleGlyAlaIleIleSerPheGlyThrLeuIleThrValPheTrpProAlaGly 120  
Db 1211 GCTTAGGTCGACGATATATATCTTTGGACTTTGGTCCGATCTTTTGG---CAAGGA 1267  
QY 121 GluGlnAspGlyThrValTrpThrGlnPheIleGlyMetGlyGluIlePheValAspThr 140  
Db 1268 TCTGACCTGCAATGTTTGGCAGAGATTTGTTAAACATCGA-----GGAAG 1315  
QY 141 ProLeuThrGlu-----SerIleGlyLeuLeuGlyLeuGlnThrLeuGluGlyPhe 157  
Db 1316 CCTATACAGAAATATGATATAAAACATATTAATGATACCTTCTATGCTACACCTATA 1375  
QY 158 ArgGlnIleLeuGlnSerTyraThrAlaLeuAspAspTrpArgLysLeuLysArgLeu 177  
Db 1376 AAAAATCAACTGATTAATATCAAGAAATTTTCGATTAATGGACGACGACACT----- 1429  
QY 178 GlnAlaProGlyLeuProProSerSerAlaLeuGlnGlnAlaIleLeuThrLeuLysIle 197  
Db 1430 -----ACACACGCTAATGCTAAAGCAGTACATGATCTTTTACT 1468  
QY 198 ArgPheGluAsnValHisAsn-----AspPheIleArgGluIleProGlyPheGln 214  
Db 1469 ACCTTAGAACCTATATATGATATAAGATTTAGATGTTAAATAAATTAATGCTAGCTATGA 1528  
QY 215 LeuGlnThrTyraLeuThrLeuLeuProIleTyraGlnAlaAlaAsnPheHisPhe 234  
Db 1529 ATACCAACA-----CTCCCTGATATGACAAATAGCTACTTGACACTTG 1573  
QY 235 AsnLeuLeuGlnGlnGlyAlaGluLeuAlaAspGluTrpAsnAlaAspIleHisProSer 254  
Db 1574 AATTATTAATAAACATGCTGCTACTACTTACATATATATG-----CTGCAAAATCA 1624  
QY 255 GlnIleGluProAsnAlaGlyThrSerAspAspTyraTyraLeu-----LeuLeuGlyGluAsn 273  
Db 1635 GGATTAATCCAGATCTTCAATCATCTATATCTATACAGGCTATTTAAACGTATA 1684  
QY 274 IleProLysTyraSerAsnTyraAlaAsnThrTyraGlyGlyLeuLysAsnLeu 293  
Db 1685 ATACAGAAATATACGACTATGTTATACAAACGTACCAATGACGACGACTATGATGTA 1744  
QY 294 AspGluProAsnMetLysTrpSerIlePheAsnAspTyraArgArgTyraMetThrIleThr 313  
Db 1745 ACTAATACTAACGCAACATGAAATATGTAATATACTTACCGTTTGAATATGACTTAACT 1804  
QY 314 ValLeuAspThrIleSerGlnPheSerLeuTyraAspIleLysArgTyraArgAspSerIle 333  
Db 1805 GGTGATGACTTATTTGCTATTTTCCAAATTAATGACCCAGAAAAATAT----- 1852  
QY 334 GlyGlyIleGluValLysGlyIleLysAsnGluLeuThrArgGluIleTyraThrGln 353  
Db 1853 -----CCAATAGAGTTAAATCTTGAACCTTATACAGAACTTAT---ACGAAT 1897  
QY 354 IleAsnPheAspArgLeuProGlnLeuArgValGlnProAsnLeuAlaThrMetGlyTyra 373  
Db 1898 GTTAATCTCAGATCACTTT-----AGAACATACAGAACTATAGAAAT 1939  
QY 374 AsnLeuThrArgAlaSerPheLysLeuPheSerPheLeuGlnPheIlePheTyraThr 393  
Db 1940 GGATTAACACTAG---AATCTCACTATTTACTTGATTAACCAAGGCGGTTTATACCA 1996  
QY 394 GluAsnThrAsn-----PheGlyAsnArgLeu 402  
Db 1997 AGAAATTCGAGACATTCCTGATCTTATGATATTTTCTTTTACAGGTAAACAGATG 2056  
QY 403 ValGlyIleSerAsnArgAspAlaProThrTyraSerAsnThrIleThrGlnThrLeuTyra 422

Db 2057 -----GCTTTACACATCTAATGATGATCGCAACATAATCTGG 2095  
QY 423 GlyGluArgThrGly-----SerProThrThrLysThrIleArgProPheGluSer 439  
Db 2096 GGAAGCGTTCATGGAATAATATTTCTTCAGACACATCCAAAGTATTTCTTTTATAGA 2155  
QY 440 Tyr-----LysValSerIleValThrAspArgGlnSerProValSerPro 455  
Db 2156 AACAAACCTATGATTAAGTCGAAATTTGTCAGACATACAGACTCTCAGATATATA--- 2212  
QY 456 IleGlnProHisPheIleIleAsnGlnIleGluLeuTyraLeuAsnGlySerSerAsnAsn 475  
Db 2213 -----TATGAATGATATTTTTTTCGAATATGACGTAA 2245  
QY 476 ThrLeuLysTyraSerAlaGlyLysLeu-----SerAsnTyraGlnAsnThrThrPhePhe 494  
Db 2246 GTATTCGATATTTATTCATCCAAATTCACAAATATAGAAATATTTATTAAGAAACGATCTTAT 2305  
QY 495 GlnPheProArgLysLysAspCysAsnLeuValIleAspProGlyCysSerProAsnPhe 514  
Db 2306 ATGATTCCAAAACAAACATGAAAAAT----- 2332  
QY 515 AsnAsnTyraSerHisIleLeuSerHisPheSerLeuPheThrTyraSerTyraIleGly 534  
Db 2333 GAAGAATATGCTCATCTATCGTATATAAAACGTAAATTAATATTTTCAGTACT 2392  
QY 535 LeuGlnLeuGlnIleLeuAspThrGlyValLeuGlyTyraProHisSerSerValAspArg 554  
Db 2393 AGAGAAAGAAAGAGAGTT-----GCATTTAGTTGACACATACATAGTGTGATTTTC 2443  
QY 555 TyraAsnAlaIleSerAspLysIleIleThrMetIleProAlaIleLysGlyAsnAsnLeu 574  
Db 2444 CAAATATCAATATGATTTGATATACATCACCAAAATCCACGCTCTAAAGCTTTGAAGGTA 2503  
QY 575 AspThrAsnSerLysValIleGluGlyProGlyHisThrGlyGlyAsnLeuValTyraLeu 594  
Db 2504 AGTCTGATGAAAAATGTAAGAGTCTGCTGACACAGGAGGAGCTTGTAATTTCTT 2563  
QY 595 GlnSerGlnGlyArgLeuGluIleThrCysGluThrProAsnSerThrGlnSerTyraPhe 614  
Db 2564 AAAGATATGATGATTTTAGAGTTAGATT---TTAAAAAATGTTTCTCGACAAATATCA 2620  
QY 615 IleArgLeuArgTyraIleThrAsnGlyAlaGlyAsnThrLeuProAsnIleSerLeuThr 634  
Db 2621 GTACGATATCGTATGCTACTAATGCT-----CCAAAGACAAACAGTATTC 2665  
QY 635 IleProGlyValIleGlyIleProProGlnArgLeuAsnAsnThrPheSerGlyThrAsn 654  
Db 2666 TTAACCGGAATATGATCTATATAGTGTGAG---CTCCCTAGTACCACTTCCGCCAAAC 2722  
QY 655 TyraAsn-----AsnLeuGlnTyraGlyAspPheGlyTyraPheGlnPheProSerThrVal 672  
Db 2723 CCAATGCTACAGATTTAAACATATGCAAGATTTTGATATGTAACCTTTCCAAAGAACGTT 2782  
QY 673 -----ThrLeuProLeuAsnArgAsnIleProPheIle 683  
Db 2783 CCAATTAACATTTGAAGAGAGAGACCTTATTAATATGACC-----TTA 2827  
QY 684 PheAsnArgAlaAspValSerAsnSerIleLeuIleIleAspLysIleGluPheIlePro 703  
Db 2828 TATGTTACACCAAAATCATATATATATA---TATATTTGCAAAATATGAAATTTATTTCCA 2884  
QY 704 IleThrSerSerMetHisGlnAsnArgGluLysGlnLysLeuGluThrIleGlnThrLys 723  
Db 2885 ATCCTCATATCTGTTTATGATTTATACAGAGAAACAAATATATAGAAAAACACAGAAAAATA 2944  
QY 724 IleAsnThrPhePheThrAsnHisIleThrLysThrLeu 735  
Db 2945 GTGAATGATTTATTTGTAATTAACAAAGTCTT 2980

RESULT 11  
BTPROTOXIS  
LOCUS BTPROTOXIS 127923 bp DNA linear BCT 16-APR-2005

```
DEFINITION Bacillus thuringiensis subsp. israelensis plasmid pBtoxis.
ACCESSION AL731825
VERSION AL731825.1 GI:21685410
KEYWORDS
ORGANISM Bacillus thuringiensis serovar israelensis
Bacillus thuringiensis serovar israelensis
Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus; Bacillus
cereus group.
REFERENCE 1 (bases 1 to 127923)
AUTHORS Berry,C., O'Neil,S., Ben-Dov,E., Jones,A.F., Murphy,L., Quail,M.A.,
Holliden,M.T., Harris,D., Zaritsky,A. and Parkhill,J.
TITLE Complete sequence and organization of pBtoxis, the toxin-coding
plasmid of Bacillus thuringiensis subsp. israelensis
JOURNAL Appl. Environ. Microbiol. 68 (10), 5082-5095 (2002)
PUBMED 12324359
REFERENCE 2 (bases 1 to 127923)
AUTHORS Parkhill,J.
TITLE Direct Submission
JOURNAL Submitted (19-APR-2002) Submitted on behalf of the pBtoxis
sequencing team, Sanger Centre, Wellcome Trust Genome Campus,
Hinxton, Cambridge CB10 1SA E-mail: parkhill@sanger.ac.uk
COMMENT
Notes:
Details of pBtoxis sequencing at the Sanger Centre are available on
the World Wide Web.
(URL, http://www.sanger.ac.uk/Projects/B_thuringiensis/).
FEATURES
source
1..127923
location/Qualifiers
/organism="Bacillus thuringiensis serovar israelensis"
/mol_type="genomic DNA"
/db_xref="taxon:1430"
/note="plasmid pBtoxis"
1145..1149
/gene="pBto01"
1163..2089
/gene="pBto01"
1163..2089
/note="pBto01"
/gene="pBto01"
/note="Similar in part to Bacillus anthracis pXOI-49
TR:Q9X319 (EMBL:AF065404) (227 aa) fasta scores: E():
8.9e-44, 78.48% id in 158 aa"
/codon_start=1
/transl_table=11
/protein_id="CAD30064.1"
/db_xref="GI:21685411"
/db_xref="UniProt/TREMBL:O8KRX5"
/translation="MSTHNFSSQNPVKHGVFPDGLNITILENERKELPSALATY
IADALVHNTAETDMINISLSYRIPFYLTQTLISLVARNDRGIMLLEIYV
FSRKIGMHVKKHIEDFETTRMAFLKEKLNNAKROYIIVPDIKFNADVLDKDK
SSSRITVRRVQVQVITEKPVIFSSNVLNKKVLEHPIAKCRKAVSRLLKMGQ
ALRKDKENIMTAFRQEIYDITDIYPTKOKN"
2246..3106
/note="IS240"
complement(2246..3106)
/rpt_type=INVERTED
2337..3044
/gene="pBto03"
2337..3044
/gene="pBto03"
2337..3044
/note="pBto03"
/note="Similar to Bacillus thuringiensis insertion element
IS240-a protein TR:Q45766 (EMBL:M23740) (235 aa) fasta
scores: E(): 2.2e-92, 99.57% id in 235 aa, and to
Mycobacterium fortuitum, transposase tnp tnp6100 TR:Q49185
(EMBL:X53635) (254 aa) fasta scores: E(): 1.4e-37, 48.05%
id in 231 aa"
/codon_start=1
/transl_table=11
/product="insertion element IS240 protein"
/protein_id="CAD30065.1"
/db_xref="GI:21685412"
/db_xref="GOA:O8KRX4"
/db_xref="InterPro:IPR001584"
/db_xref="UniProt/TREMBL:O8KRX4"
/translation="MEKENIFPKKHVYQADMILMTVRYLRYNLSPRDLVEMMERGLS
misc_feature
2538..2933
/gene="pBto03"
/note="HMPFfam hit to PF00665, Integrase core domain"
3090..3106
/rpt_type=INVERTED
3521..4381
/note="IS240"
complement(3521..4381)
/rpt_type=INVERTED
3612..4319
/gene="pBto04"
3612..4319
/gene="pBto04"
/note="pBto04"
/note="Similar to Bacillus thuringiensis insertion element
IS240-a protein TR:Q45766 (EMBL:M23740) (235 aa) fasta
scores: E(): 3.5e-91, 99.14% id in 235 aa, and to
Mycobacterium fortuitum, transposase tnpA or tnp6100
TR:Q49185 (EMBL:X53635) (254 aa) fasta scores: E():
1.1e-37, 48.05% id in 231 aa"
/codon_start=1
/transl_table=11
/product="insertion element IS240 protein"
/protein_id="CAD30066.1"
/db_xref="GI:21685413"
/db_xref="GOA:O8KX5"
/db_xref="InterPro:IPR001584"
/db_xref="UniProt/TREMBL:O8KX5"
/translation="MEKENIFPKKHVYQADMILMTVRYLRYNLSPRDLVEMMERGLS
LSHTTMRVHGYGPELNERIRKHLKRTNDSNRVDETYIKIGENMYLRAVDSGNT
LDPLYSKRDAKAAKCPFKKALASPVHTKPRVITVDGNKAYVVALRELKNSIPYGM
PLRVKVKYLNMMIEDHPRFKKRILNMLGKSKQTVKNIAGIEAHMVKKGLKLRQ
SAQONRCIHQFGLTA"
3813..4208
/gene="pBto04"
/note="HMPFfam hit to PF00665, Integrase core domain"
4365..4381
/rpt_type=INVERTED
complement(4365..4381)
/gene="pBto05"
complement(4634..5275)
/gene="pBto05"
/note="Similar to N-terminus of Bacillus anthracis pXOI-18
TR:Q9X2Y9 (EMBL:AF065404) (315 aa) fasta scores: E():
4.2e-56, 84.15% id in 183 aa, and to Bacillus halodurans
Bh2364 protein TR:Q9KAC5 (EMBL:AP001515) (378 aa) fasta
scores: E(): 1.6e-18, 35.45% id in 189 aa, and weakly to
Lactobacillus delbrueckii integrase/recombinase orf2
TR:Q48538 (EMBL:Z50864) (333 aa) fasta scores: E(): 6.3,
28.88% id in 90 aa, and to Bacillus thuringiensis
resolvase tnpI SW:TNRI BACTU (P10020) (284 aa) fasta
scores: E(): 8.5, 23.88% id in 180 aa"
/codon_start=1
/transl_table=11
/product="Integrase/recombinase family protein"
/protein_id="CAD30067.1"
/db_xref="GI:21685414"
/db_xref="GOA:O8KRX3"
/db_xref="InterPro:IPR004107"
/db_xref="UniProt/TREMBL:O8KRX3"
/translation="MSKNINLEETLAASAVYIEKGRQSTIKRYAYDIKDPYKMLRA
NKKLIHKSWSELSEADYQTFSELEDRKYSQKTRHILWVLKGLHMFICVSPIDG
INLSLIPDQSLNDNDPITEMEKKILKOTVLSTKGLTEQAKXRPILMDNVCITLNV
NYGSLQGLVLSNHSNINPARINIMCYVKKIDTIANDIIFVQGLAVAFHRTSF"
complement(4985..5251)
/gene="pBto05"
/note="HMPFfam hit to PF02899, Phage integrase, N-terminal
SAM-like domain"
complement(5283..5286)
complement(5364..5507)
```



```

Db 40117 AGAATTCGAGACATTCCTGATCTTATGATATTTTCTTTTACAGTAAACGATG 40176
QY 403 ValGylIleSerAsnArgAspAlaProThrTySerAsnThrIleThrdIuThrLeuYr 422
Db 40177 -----GCCCTTACACATCTACTAATGATGATCGCAACATATATCGG 40215
QY 423 GlyIuArgThrGly-----SerProThrThrTySerIleArgProPheGluSer 439
Db 40216 GGAAGCGGTTCATGAAATATTTTCTCTCAAGACATCCAAAGTATTTCTTTATAGA 40275
QY 440 Tyr-----LysValSerIleValThrAspArgGlnSerProValSerPro 455
Db 40276 AACAAACCTATGATGATGATGCAAAATGTCACATAGAGAGTACTCATGATATATA--- 40332
QY 456 IleGlnProHisPheIleIleAsnGlnIleGluLeuTyLeuAsnGlySerSerAsnAsn 475
Db 40333 -----TATGAAATGATATATTTTTCGAAATGCGACGTGA 40365
QY 476 ThrLeuTySerIleGlyGlySerLeu---SerAsnTyLeuAsnThrThrPhePhe 494
Db 40366 GTATTTCGATATTCATCCATTCACATGAAATATATATATAAGATCTGATTCCTAT 40425
QY 495 GlnPheProArgIleLysAspCysAsnLeuValIleAspProGlyCysSerProAsnPhe 514
Db 40426 ATGATTCGAAACAAACATGCAAAAT----- 40452
QY 515 AsnAsnTySerHisIleLeuSerHisPheSerLeuPheThrTySerTyValIleGly 534
Db 40453 AAAGATATGTCATACATCTATCTGATATATAAACTGATATATATATTTTCACTAGT 40512
QY 555 LeuGlnLeuGlnIleLeuAspThrGlyValLeuGlyTTPThrHisSerSerValAspArg 554
Db 40513 AGAGAAAGAGAGAGTT-----GCATTAGTTGACACATACATAGTTTATATTC 40563
QY 555 TyrAsnAlaIleSerAspIleIleThrMetIleProAlaIleLysGlyAsnAsnLeu 574
Db 40564 CAAATATACATGATTTAGATACATCACCAATCCACGCTCTTAAACCTTGAAGTA 40623
QY 575 AspThrAsnSerLysValIleGlnGlyProGlyHisThrGlyGlyAsnLeuValTyLeu 594
Db 40624 AGTTCGATTCGAAATATGTCGAAAGTCTCGTGTACACAGGTGAGAGACTGTGTAATCTT 40683
QY 555 GlnSerGlnGlyArgLeuGlnIleThrCysGlnIuThrProAsnSerThrGlnSerTyPhe 614
Db 40684 AAAGATAGTATGATTTTATGAGTTGATTT--TTAAAAATGTTTCTCGACATATACA 40740
QY 615 IleArgLeuArgTyAlaThrAsnGlyAlaGlyAsnThrLeuProAsnIleSerLeuThr 634
Db 40741 GTACGATATTCGTTATGCTACTATAGCT-----CMAAGACAAACAGTATTC 40785
QY 635 IleProGlyValIleGlyIleProProGlnArgLeuAsnAsnThrPheSerGlyThrAsn 654
Db 40786 TTAACCGGAATGATGATCTATAGTGTGAG--CTCCCTAGTACCACTTCGCCCAAAAC 40842
QY 655 TyrAsn-----AsnLeuGlnIuTyArgLysPheGlyTyPheGlnPheProSerThrVal 672
Db 40843 CCAATGCTACAGATTTTAAATATGACAGTTCGATTTTGAATATGTAACATTTCCAGAAACAGT 40902
QY 673 -----ThrLeuProLeuAsnArgAsnIleProHelle 683
Db 40903 CCAATTAATAACATTGGAAGAGACACACTTATATATGAC-----TAA 40947
QY 684 PheAsnArgAlaAspValSerAsnSerIleLeuIleIleAspLysIleGluPheIlePro 703
Db 40948 TATGTGACACCAATCATTCATATATAATA---TATATTGACAAATATGAAATTTATTCAC 41004
QY 704 IleThrSerSerMetHisGlnAsnArgGlyGlyGlnLysLeuGlnThrIleGlnThrLys 723
Db 41005 ATCACTTCATCTGTATGATATATACAGAGAGCAAAATATATAGAAAAAACACAGAAATA 41064
QY 724 IleAsnThrPhePheThrAsnHisThrLysThrLeu 735
Db 41065 GTGAATGATTTATTTGTTAATTAACAAAGTTCTT 41100

```

```

RESULT 12
E00614 3756 bp DNA linear PAT 29-SEP-1997
LOCUS E00614
DEFINITION DNA encoding a polypeptide having insecticidal activity (Bt
endotoxin).
ACCESSION E00614
VERSION E00614.1 GI:2168893
KEYWORDS JP 1986005098-A/1.
SOURCE Bacillus thuringiensis
ORGANISM Bacillus thuringiensis
Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus; Bacillus
cereus group.
1 (bases 1 to 3756)
REFERENCE
AUTHORS Aran, R. U. and Toomaru, J. P.
TITLE INSECTICIDAL POLYPEPTIDE
JOURNAL Patent: JP 1986005098-A 1 10-JAN-1986;
SHINTORO CORP
OS Bacillus thuringiensis
PN JP 1986005098-A/1
PD 10-JAN-1986
PR 21-FEB-1985 JP 1985031647
PR 22-FEB-1984 US 84 582506, 22-JAN-1985 US 85 633556 PT
ARAN EMU UORUFUIRUDO, TOOMASU JIEI POROTSUKU PC
C07K15/04, A01N63/02, C07H21/04, C12N15/00//C12P21/00, (C12N15/00, PC
C12R1:19);
CC screenedness: Single;
CC topology: Linear;
CC hypothetical: No;
CC anti-sense: No;
CC *source: straln=BtRa4-1;
CC *source: clone=BtRa4-1;
FH key Location/Qualifiers
FH 5'UTR <1..925
FT CDS 926..2971
FT FT /product='Bt endotoxin'
FT mat_peptide 926..2968
FT FT /product='Bt endotoxin'
FT 3'UTR 2972..3756.
Location/Qualifiers
1..3756
/organism='Bacillus thuringiensis'
/mol_type='genomic DNA'
/db_xref='taxon:1428'

```

## ORIGIN

```

Alignment Scores:
Pred. No.: 1.49e-56 Length: 3756
Score: 880.00 Matches: 245
Percent Similarity: 47.9% Conservative: 123
Best Local Similarity: 31.9% Mismatches: 279
Query Match: 22.7% Indels: 122
DB: 6 Gaps: 31

```

US-10-783-417-2 (1-735) x E00614 (1-3756)

```

QY 1 MetAsnGlnAsnAspAsnAsnGluTyrgluIleIleAspSerHisThrSerProTyR 20
Db 941 ATGAATCCATATCAAAATATGAAATGATATGAAATATTCATGCTCATCCATGCTTTT 1000
QY 21 PheProAsnArgAsnSerAsnAspSerArgTyProTyrThrAsnAsnProAsnGlnPro 40
Db 1001 -----ACCAAGTCTAATATGATATTCATGATATTCATTAAGCAAAATAGCCAAATCAACCA 1054
QY 41 LeuGlnAsnThrAsnTyLysGluTyLeuAsnMetCysGlnGlyAsnThrGlnTyrgly 60
Db 1055 CTGAATAACACGAATATGAAAGATGGCTCAATGTGTGTCAGATTAATCAACAAATATGCC 1114
QY 61 AspAsnPheGluThrPheAlaSerAlaAspThrIleAlaAlaValSerAlaGlyThrIle 80
Db 1115 AATATGCGGGGAATTTTGTCTAGTTCGAACTATGTTGGAGGTGATGAGGTATATAT 1174

```

QY 81 ValSerGlyThrLeuLeuAlaGlyIleGlyLeuThrSerIleSerGlyProIleGly 100  
 Db 1175 GTAGTAGAAGAACTATGTAGA-----GCTTTGCTGCCCT----- 1210  
 QY 101 IleIleGlyAlaIleIleSerPheGlyThrLeuIleThrValPheTrpProAlaGly 120  
 Db 1211 GCTTAGCTGCGAGATATATCTTTTGGACTTTTGGCTGCGATCTTTGG---CAAGA 1267  
 QY 121 GluGlnAspLysThrValTrpThrGlnPheIleLysMetGlyGluIlePheValAspThr 140  
 Db 1268 TCTGACCCCTGCAAAATGTTTGGACAGATTTGTAACATCGA-----GGAAGG 1315  
 QY 141 ProLeuThrGlu-----SerIleGlyGlnLysLeuGlnThrLeuGlnGlyPhe 157  
 Db 1316 CCTATACAGAAATAGATATAAAACATATTAATTAATGTAACCTTCTATGACACTATA 1375  
 QY 158 ArgGlnIleLeuGlnSerThrAsnThrAlaLeuAspAspTrpArgLysLeuLysArgLeu 177  
 Db 1376 AAAAATCAACTGATTAATATCAAGAAATTTTTCGATTAATGAGACGACACGCT----- 1429  
 QY 178 GlnAlaProGlyLeuProProSerSerAlaLeuGlnGlnAlaAlaLeuThrLysIle 197  
 Db 1430 -----ACACACGCTATGCTAAAGCAGTACGATCTCTTACT 1468  
 QY 198 ArgPheGluAsnValHisAsn-----AspPheIleArgGluIleProGlyPheGln 214  
 Db 1469 ACCTTAGAACCTATATAGATTAAGATTTAGATGTTAAAAAATATAGCTAGCTATGCA 1528  
 QY 215 LeuGlnThrTrpLysThrLeuLeuLeuProIleTrpAlaGlnAlaAlaAsnPheHisLeu 234  
 Db 1529 ATACCAACA-----CTCCCTGCATATGACAAATAGCTATCTGGACATCG 1573  
 QY 235 AsnLeuLeuGlnGlnGlyValaGluLeuAlaAspGluTrpAsnAlaAspIleHisProSer 254  
 Db 1574 AATTATATTAACACATGCTGCTACCTATTAACATATATG-----CTGCAAAATCA 1624  
 QY 255 GlnIleGluProAsnAlaGlyThrSerAspAspTrpTrpLys--LeuLeuLysGluAsn 273  
 Db 1625 GGTATTAATCCAAAGTACTTCAATCATCTAATTAATCAAGGCTATTTAAACGTTAA 1684  
 QY 274 IleProLysTrpSerAsnThrCysAlaAsnThrTrpArgThrGlyLeuLysAsnLeuArg 293  
 Db 1685 ATACAGAAATATACGACTATGTATACAAACGTAACATGACGACGTAATGATGTA 1744  
 QY 294 AspGluProAsnMetLysTrpSerIlePheAsnAspTrpArgArgTrpMetThrIleThr 313  
 Db 1745 ACTATATACTAACGCAACATGGAATATGATATATCTACCGTTTGAATGACTTAACT 1804  
 QY 314 ValLeuAspThrIleSerGlnPheSerLeuTrpAspIleLysArgTrpArgAspSerIle 333  
 Db 1805 GTGTATAGATCTTATGCTATTTTCCAAATTAATGACCCAGAAAATAT----- 1852  
 QY 334 GlyGlyIleGluValLysGlyIleLysAsnGluLeuThrArgGluIleTrpThrGlu 353  
 Db 1853 -----CCAATAGAGTTAAATCTGAACCTTACAGGAAGTTAT--ACGAAT 1897  
 QY 354 IleAsnPheAspArgLeuProGlnLeuArgValGlnProAsnLeuAlaThrMetGluTrp 373  
 Db 1898 GTTATATGAGTACATTT-----AGAACATACAGAACTAATAAAT 1939  
 QY 374 AsnLeuThrArgAlaSerPheLysLeuPheSerPheLeuGlnGlnPheIlePheTrpThr 393  
 Db 1940 GGATTAACATAGA--AATCTACATTATTTACTTGATTAACCAAGGGGGTTTTTACACA 1996  
 QY 394 GluAsnThrAsn-----PheGlyAsnArgLeu 402  
 Db 1997 AGAAATTTCTCGACACATTTCTGATCTTAATGATATTTTCTTTTACAGGTAAACAGATG 2056  
 QY 403 ValGlyIleSerAsnArgAspAlaProThrTrpSerAsnThrIleThrGluThrLeuTrp 422  
 Db 2057 GCCTTACACATACATTAATGATAT-----CGCACATATATCTGGACGACGTTCAAT 2107  
 QY 423 GlyLysArgThrGlySerProThrThrLysThrIle-----ArgProPhe 437

Db 2108 GGAATATATTATTTCTCAAGACACATCCAAAGATTTTCTTTTATAGAAACAAACTATTT 2167  
 QY 438 GluSerTrpLysValSerIleValThrAspArgGlnSerProProValSerProIleGln 457  
 Db 2168 GAT-----AGGTCGAATTTGTCAGACATAGAGATCTCGATATATA----- 2212  
 QY 458 ProHisPheIleIleAsnGlnIleGluLeuTrpLeuAsnGlySerSerAsnAsnThrLeu 477  
 Db 2213 -----TATGAATGATATTTTTCGAAATATGACGTAAGATATTT 2251  
 QY 478 LysTrpSerAlaGlyGlySerLeu--SerLeuTrpGlnAsnThrPhePheGlnPhe 496  
 Db 2252 CGATATTCATCCAAATTCACAAATAGAAATATATTAATAAAGAACTGATTTCTTATATGATT 2311  
 QY 497 ProArgLysLysAspCysAsnLeuValIleAspProGlyCysSerProAsnPheAsnAsn 516  
 Db 2312 CCAAAACAAATGGAATAAT-----GAAAGA 2338  
 QY 517 TrpSerHisIleLeuSerHisPheSerLeuPheThrTrpSerTrpValIleGlyLeuGln 536  
 Db 2339 TATGTCATACCTATGCTATATATAAAGCTATATATATATATTTTCAGTATGAGAA 2398  
 QY 537 LeuGlnIleLeuAspThrGlyValLeuGlyTrpThrHisSerSerValAspArgTrpAsn 556  
 Db 2399 AGAAGAAAGATT-----GCATTTAGTTGGACACATCTAGTGTGATTTCCAAAT 2449  
 QY 557 AlaIleSerAspLysIleIleThrMetIleProAlaIleLysGlyAsnAsnLeuAspThr 576  
 Db 2450 ACAATAGTTTATAGTATACATACCAATCCAAATCCAGCTCTAAATACTTTAAAGTAATCT 2509  
 QY 577 AsnSerLysValIleGluGlyProGlyValIleThrGlyGlyAsnLeuValTrpLeuGlnSer 596  
 Db 2510 AATCAAAAATTTGGAAGGCGCTCGCTCACAAAGTAAAACTTGTAATCTTATAAGAT 2569  
 QY 597 GlnGlyArgLeuGluIleThrCysGluThrProAsnSerThrGlnSerTrpPheIleArg 616  
 Db 2570 AGTATGAATTTTAAAGTTAGATT--TTAAAAATGTTTCTCAACAAATATCAAGTACGT 2626  
 QY 617 LeuArgTrpAlaThrAsnGlyValaGlyAsnThrLeuProAsnIleSerLeuThrIlePro 636  
 Db 2627 ATTGTTATGCTACTATATGCT-----CCAAAGACACAGTATCTTAAC 2671  
 QY 637 GlyValIleGlyIleProProGlnArgLeuAsnAsnThrPheSerGlyThrAsnTrpAsn 656  
 Db 2672 GGAATAGTACTATTAAGTGTGAG--CTCCCTAGTACCACTCCGCGCAAAACCCAAAT 2728  
 QY 657 -----AsnLeuGlnTrpGlyLysAspPheGlyTrpPheGlnPheProSerThrValThrLeu 674  
 Db 2729 GCTACAGATTTTAAACATATGACAGATTTTGGATATGTAATTTTCAMGAACGTTTCA-- 2785  
 QY 675 ProLeuAsnArgAsnIle-----ProPheIle-----PheAsnArg 686  
 Db 2786 -----AATTAATAACATTTGAAGAGAGACACTTTATTAATGACCTTTATATGTGTCA 2839  
 QY 687 AlaAspValSerAsnSerIleLeuIleIleAspLysIleGluPheIleProIleThrSer 706  
 Db 2840 CCAATATCTCATATATATATA--TATATGACAAATATGATATTTATTCATCACTCA 2896  
 QY 707 SerPheHisGlnAsnArgGluLysGlnLysLeuGluThrIleGlnThrLysIleAsnThr 726  
 Db 2897 TCTGATATAGATTATTAACAGAAAGCAAAATATAGAAAAACACAGAAATATGATGATAT 2956  
 QY 727 PhePheThrAsnHisThrLysThrLeu 735  
 Db 2957 TTAATTTGTAATTAATAACAAAGTTCTT 2983  
 RESULT 13  
 AF285775 4959 bp DNA linear BCT 16-AUG-2000  
 LOCUS Bacillus thuringiensis subsp. finitimus parasporal inclusion  
 DEFINITION protein Cry (cry) gene, complete cds.  
 ACCESSION AF285775

VERSION AF285775.1 GI:9828609  
KEYWORDS  
SOURCE Bacillus thuringiensis serovar finitimus  
ORGANISM Bacillus thuringiensis serovar finitimus  
Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus; Bacillus cereus group.  
REFERENCE  
AUTHORS Moore, G.D. and Debro, L.H.  
TITLE Cloning and Analysis of a Gene for Exoprotein Enclosed Cry Protein of Bacillus thuringiensis subspecies finitimus  
JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 4959)  
AUTHORS Moore, G.D. and Debro, L.H.  
TITLE Direct Submission  
JOURNAL Submitted (07-JUL-2000) Biology, Jacksonville State University, 700 N. Pelham Road, Jacksonville, AL 36265, USA  
FEATURES  
source  
1..4959  
/organism="Bacillus thuringiensis serovar finitimus"  
/mol\_type="genomic DNA"  
/sub\_species="finitimus"  
/db\_xref="taxon:29337"  
1128..4514  
/gene="cry"  
1128..4514  
/gene="cry"  
/codon\_start=1  
/transl\_table=11  
/product="parasporal inclusion protein Cry"  
/protein\_id="AA00235.1"  
/db\_xref="GI:9828610"  
translation="MDKHINKIGVSTEVNSRIFPNRVDSDTVAVASGIVVGT  
ILTAASVFNRGVULISGCTLAPVWPEBEPKRIQGMKRGDILNQITSTAVKE  
IALAHUNGKDVLTIERAFNDKRNPSANTARLVSOENAHFVSNMQLQPT  
DTLLSCYTEANMLHLNHOQVADQPMNDQPSPMKSGTYDDELVIETKYN  
YCTKTHGKLNHMKESKRTWAVNTYRREMTLLVLDVATPPYDIRPFGVLEL  
TRVYTSLDPPGLNAPLPEIDPSYIEHLETPGLFVLSIDIELTYSVAGDVLSG  
IRSKYVGNQFTMKNIYGNMRLSKDLITLPGRIITHLSINGPOTIAGINKLS  
LIQKIVPTFRKNDNEYQKRNVNONEQETTNYPNDVGSNSQPKNLSHPLIIT  
QVEFAVPHSIFALGWTNNSVNSQNLISESVTOIPLVAHEVTNVSIRGFGSD  
LIELNDKSIKCKKASLKKYALSLFYAANMALAVSIDVDSAGVTLQPTSRKNNN  
FTIOLNATKDFQHTLLVDIEIPSESEIHIHLKRDDEYEGVILLIDLEFPYIDENY  
TNEMNLKAKKAVNVLFINATVALKMDVDYHIDQVAMVEGIDSLAKKIKLHA  
IKKAKOLSOAVRLSDPNPNMNAENSGWTGVTITGDPYKGRALIOASAAPNEF  
PTLYXOKIDESLAKPYTRYQLRGVTEGSDLELDVIRGATDITVAVPQDELTISYA  
PINPCEIETRLDITTCALDRCKSNVNSADVAPDDVGGPHAFSHLIDTDDN  
RNLGIMILFKIATPDGYATFGNLELIEGLPSGALAQVORKEQMGKNTQKREAA  
KLVAAKOTINQLPADSQGKLRPTEFSNLSADKLVIKIDVYSVLSVPLGNYD  
LMELENRIONALIDLVDARNVTYNGEFGNLANMAASNTTEVROIQDTSVLSSMNA  
OVAOSIANRPHGVTLYRTAKKEGICNGCVITLDCANHIDPLITPSSCDSGPTSSNEL  
AAVYTKLIEIPFDIQIRIEIGTEGTYYVESVDLIRMED"

ORIGIN  
Alignment Scores:  
Pred. No.: 4,466-55 Length: 4959  
Score: 862.50 Matches: 234  
Percent Similarity: 50.1% Conservative: 112  
Best Local Similarity: 33.9% Mismatches: 252  
Query Match: 22.3% Indels: 93  
DB: 1 Gaps: 27  
US-10-783-417-2 (1-735) x AF285775 (1-4959)

QY 68 SerAlaapThrTrlAlaAlaValSerAlaGlyThrTrlLevalSerGlyThrLeuLeuAla 87  
DB 1209 AGCACTGATACAGCTGCTAGTACCGCAGGATAGTGTAGTGGTACTACTACCA 1268  
QY 88 GlyIleGlyGlyLeuThrSerTrlSerGlyProIleGlyIleIleGlyAlaIleIleIle 107  
DB 1269 -----GCCTTGCATCATTTGTTAAATCA-----GGTGTGGTACTATA 1307  
QY 108 SerPheGlyThrLeuIleThrValPheTrpProAlaGlyGluGluAsp---LysThrVal 126  
|||||

DB 1308 TCATTGGACCTTGGCTCCCGTTCTTGGCCCTGATCCAGAGAAATCCAAAAAATT 1367  
QY TrpThrGlnPheIleValMetGlyGluIlePheValAspThrProLeuThrGluSerIle 146  
DB 1368 TGGTACAAATTATATAACACGAGAGAACCTTTTAAATCAAAATTTTACAGCTGTA 1427  
QY LysGlnLeuLysLeuGlnThrLeuGluGlyPheArgGlnIleLeuGlnSerTrpAsnThr 166  
DB 1428 AAGAAATAGCATTTAGCTCATCTTAATGTTTAAAGATGATTTAACGACTATGAAGA 1487  
QY AlaLeuAspAspTrpArgLysLeuLysArgLeuGlnAlaProGlyLeuProProSerSer 186  
DB 1488 GCATTAAATGATTTGAGAGAGA-----AATCCAAATGCA 1520  
QY AlaLeuGlnGlnAlaAlaLeuThrLeuLysIleArgPheGluAsnValHisAsnAspPhe 206  
DB 1521 -----AATCTGCCAGATTGGTATACACAGAAATTTGAAAGCCTCATTTCAATTT 1571  
QY IleArgGluIleProGlyPheGlnLeuGluThrTrpLysThrLeuLeuProIleTrp 226  
DB 1572 GTAAACCATATGCCCAACACTCCACTCCAGTATGACACTTATTATTAGTGTCTAT 1631  
QY AlaGlnAlaAlaAsnPheHisLeuAsnLeuLeuGlnGlnGlyAlaGluLeuAlaAspGlu 246  
DB 1632 ACAGAGCTGCAGAAATTTACATTTGAATTTATACATCAAGGTGTCAATTCGCGATCA 1691  
QY TrpAsnAlaAspTrlLeHisProSerGlnIleGluProAsnAlaGlyThrSerAspAspTrp 266  
DB 1692 TGGAAATGACAGATCAACCACTTACCAATGTTGAAGTATCAGTACT-----TAT 1742  
QY TyrLysLeuLeuLysGluAsnIleProLysTrpSerAsnTrpCysAlaAsnThrTrpArg 286  
DB 1743 TATGACAGCTTATTTGCTATATTTGCAAAAGTATTAATTATTGACCAAGACATACAT 1802  
QY ThrGlyLeuLysAsnLeuArgAspGluProAsnMetLysTrpSerIlePheAsnAspTrp 306  
DB 1803 AAGGATTTGAATCACTTAAAGAAATCAGAAAAATCAGATGGAGCTTATTAACATAT 1862  
QY ArgArgTrpMetThrIleThrValLeuAspThrIleSerGlnPheSerLeuTrpAspIle 326  
DB 1863 CGTCAGAAAGACCTTAATGTATGTATGATCTTGCAGCACTTTCTTTTATATATATA 1922  
QY LysArgTrpArgAspSerIleGlyLysIleGluValLysGlyIleLysAsnGluLeuThr 346  
DB 1923 GGTGCTTT-----CCAGAGGAGTGAGAACTAGAAATTAACA 1958  
QY ArgGluIleTrpThr-----GluLeuAsnPheAspArgLeuProGluLeu 362  
DB 1959 AAGAGGTTTATACAAAGTTTATGATTCACCTGCTTAAATGCTGGCCCAATACAGAAATA 2018  
QY ArgValGlnProAsnLeuAlaThrMetGluTrpAsnLeuThrArgAlaSerPheLysLeu 382  
DB 2019 -----GATTCAGTTATTTGAGAGATCACTTTAACACGACCA-----GGGCTA 2063  
QY PheSerPheLeuGluGlnPheIlePheTrpThrGluAsnThrAsnPheLysAsnArgLeu 402  
DB 2064 TTTACTTGCTGCTGCAGATTTAGCTTATACACGAGAGCTGCGCAAGCGCATTTATTA 2123  
QY ValGlyIleSerAsnArgAspAlaProThrTrpSerAsnThr-----IleThrGlu 419  
DB 2124 TCAGGTAT-----CGAGAGCTTAATATATATATATATCAATTTTTCAGATGAAA 2177  
QY ThrLeuTrpGlyGluArgThrGly---SerProThrTrpLysThrIleArgProPheGlu 438  
DB 2178 AATATTATGTGTAATACAAATAGATTAAGTAAAGCAGCTCATTAACCTTTATTCACAGCGGA 2237  
QY SerTrpLysValSerIleValThrAspArgGlnSerProProValSerProIleGlnPro 458  
DB 2238 ---TTATACCTCACTTAAGCATTAACCGCTGTTTCAAAACATATGCTGCTGTTAATAAG 2294  
QY HisPheIleIleAsnGlnIleGluLeuTrpLysAsnGlySerSerAsnAsnThrLeuLys 478  
DB 2295 TTATACGTTTAATTCAAAAAATGATTTATCAACTTTTAAAAAAGATAT-----GAA 2348

QY	479	TYRSERIALAGLYLSErLeuSerLeuSerLeuTy	-----GlnAenThrThrPhePheGln	495
Db	2349	TATCAAAAAAATTTTAAATGTGATATATCAAAATGAACCTCAAGAAACTACAA	-----	2399
QY	496	PheProArgLyLyLeuAspCySAenLeuValIleAspProGlyCySeSerProAsnPhaen	515	
Db	2400	-----	AACTATCTTAATGATTTATGTGGTGTCA---AACGCCAA	2438
QY	516	AenTySerHisIleLeuSerHisPheSerLeuPheThrTySerTyValIleGlyLeu	535	
Db	2436	AAATTCAAAACAAATTAATTTATCTCATTTCCATTAATCAATCCAC-	-----	2477
QY	536	GlnLeuGlnIleLeuAsp-----ThrGlyValLeuGlyTrpThrHisSerSer	551	
Db	2478	CAAGTAAGAGTGTCTGAGTATTTCACCTCTATATTGGCATTAAGTTGACACACAAATG	2537	
QY	552	ValAspArgTyAspAlaIleSerAspLySileIleThrMetIleProAlaIleGly	571	
Db	2538	GTAACCTCCCAAAATTTATATACAAAAGTGAAGTACACAAATCCATGGTAAACCT	2597	
QY	572	AsnAsnLeuAspThrAsnSerLyValIleGlyLysProGlyHisIleThrGlyAsnLeu	591	
Db	2598	TACGAAGTACACCAATTCACCTATTAAGACACACAGGTTTAAAGGTGAAGTAAATTA	2654	
QY	592	ValTyLeuLeuSerGlnGlyArgLeuGlnIleThrCysGluThrProAsnThrGln	611	
Db	2655	ATAGAACCTTCGGAT-----AAATGTTCTATTAAATGTAAGAGCT--AGTCTTTAA	2705	
QY	612	SerTyPheIleArgLeuArgTyAlaThrAsnGlyAlaGlyAsnThrLeuProAsnIle	631	
Db	2706	AAATACGCTATAGCTATATTATATGCTGCAAAATAACGAATAGCTATCAATACAGTA	2765	
QY	632	SerLeuThrIleProGlyValIleGlyIleProProGlnArgLeuAsnAsnThrPheSer	651	
Db	2766	GGTGAATCCGACGACGAGGTCTCA-----TTCACACTTACCTTTTCT	2807	
QY	652	GlyThrAsnTyAsnAsn-----LeuGlnTyArgLysPheGlyTyPhe	666	
Db	2808	AGAAAAGGACAAATTAATTTTCAATTCACACCTTAACATAAGATTTTCAATATCAT	2867	
QY	667	GlnPheProSerThrValThrLeuProLeuAsnArgAsnIleProPheIlePheAsnArg	686	
Db	2868	ACACTTTAGTATGATTAATGAATTACCCGAAGTGAAGAAATTCATATCCATTTGAAGCA	2927	
QY	687	AlaAsp-----ValSerAsnSerIleLeuIleIleAspLySileGluPheIleProIle	704	
Db	2928	GAGATGATTAAGACGAGAGAGTGAATCTTTTAATGATTAATTAATGAGTCAAACTATA	2987	
QY	705	ThrSerSerMetHisGlnAsnArgGluLysGlnLysLeuGluThrIleGlnIleTySile	724	
Db	2988	GATGAAAATTA-----ACTAATGAAGAATTTTAGAGAGCAAGCAAGAAAGCAATG	3038	
QY	725	AenThrPhePheThrAsnHisThrLySerIleu	735	
Db	3039	AATGTATTATTATTAACGCAACAACAGCTTG	3071	

RESULT 14  
 BTH251977 6009 bp DNA linear BCT 22-DEC-2001  
 LOCUS Bacillus thuringiensis subsp. medellin cry29Aa gene for Cry29Aa  
 DEFINITION protein.  
 ACCESSION AJ251977  
 VERSION AJ251977.1 GI:17977978  
 KEYWORDS cry29Aa gene; Cry29Aa protein.  
 SOURCE Bacillus thuringiensis serovar medellin  
 ORGANISM Bacillus thuringiensis serovar medellin  
 Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus; Bacillus cereus group.

REFERENCE  
 1 Delecluse, A. and Orduz, S.  
 Characterization of two new mosquitocidal toxins, Cry29A and Cry30A, from Bacillus thuringiensis medellin

JOURNAL Unpublished  
 REFERENCE 2 (bases 1 to 6009)  
 AUTHORS Delecluse A.  
 TITLE Direct Submission  
 JOURNAL Submitted (22-Dec-1999) Delecluse A., Bacteries & Champignons  
 Entomopathogenes, Institut Pasteur, 25, rue du Dr Roux, 75724 Paris  
 Cedex 15, FRANCE

FEATURES  
 source Location/Qualifiers  
 1..6009  
 /organism="Bacillus thuringiensis serovar medellin"  
 /mol\_type="genomic DNA"  
 /isolate="163-131"  
 /sub\_species="medellin"  
 /db\_xref="taxon:79672"  
 /country="Colombia"  
 642..2594  
 /gene="cry29Aa"  
 642..2594  
 /gene="cry29Aa"  
 /function="putative moegultocidal toxin"  
 /codon\_start=1  
 /evidence=experimental  
 /transl\_table=1  
 /product="Cry29Aa protein"  
 /protein\_id="CA80985.1"  
 /db\_xref="GI:17977979"  
 /db\_xref="GOA:Q8VNX2"  
 /db\_xref="UniProt/TREMBL:Q8VNX2"  
 /translation="MNGFQNKNEYELINAPSSLSINPNYSKPIANTLNTQNTONTY  
 KDMNCAIDNNLKSINPFANLONSIVGLFALTAALASLSAPITGGSTIAAGTALA  
 AAIPIPIPMPOENNLPKKLAISEATYSRODQRYEPDALTRLSIKDSYKPEMPT  
 FMINPSTINTVTRERFQSVNGFPGSMAPFPAKATPEPILSTYQAARLHLIHLSD  
 GITYAEKMSLSRQDDMPGDLKKEPKYCNERYIEHCITKYNSSLILSKSVGANWLEY  
 NOYETPLASVLDVLSFSSYDPLRYERLSVILTRKLYTDPINHYRGISLADSEK  
 YLTPETLPTOLYTLTFPSNIFVNYMGHTNTYRLVSPKIPABRSFGKSGSYIDKVPVY  
 PINKSLIYKIRAYDNHNLFNWYFGFMDRKOIOIKIGSSRTIYIKRCHTRIAADV  
 ISHDLERKNCYSFPAWTSTTISLENEIKNDITQIPAYKAYQUGQSOVYKGGHGTGG  
 DLINIKSNDYLRISCOHLSVTKYFPRIRIVATNGSLNTRPIINITIPGTPQGMVD  
 NTEGTSQSNLYEONFGYKEFLKEVTLNPNQGISLTILNRDONSITLLDRLEFLPI  
 TPSIRSRKEQONLEKQNTVKKLFNN"

ORIGIN

Alignment Scores:  
 Pred. No.: 5.01e-54 Length: 6009  
 Score: 850.00 Matches: 232  
 Percent Similarity: 47.0% Conservative: 124  
 Best local Similarity: 30.6% Mismatches: 276  
 Query Match: 22.0% Indels: 126  
 Db: 1 Gaps: 22

US-10-783-417-2 (1-735) x BTH251977 (1-6009)

OY 1 MetAaGlnAaAaAaAaAaAaGlnGluTyrGluIleIleAaSPSerHisThrSerProTyr 20  
 Db 642 ATGAAATTCATTCAAAAATAAAAAAGAAATATGAATAATTAATATGTCATGAGT----- 695  
 OY 21 PheProAaAaAaAaAaAaAaAaAaSerAaAaAaSerAaAaAaProTyrThrAaAaAaProAaGlnPro 40  
 Db 696 ATTCTTAATATACTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 755  
 OY 41 LeuGlnAa 58  
 Db 756 ATGCAAAAACCAATTCACAAAGATTGGAAGTAAATATGTGTGCAATATGATTAATTAATTAATTAATTA 815  
 OY 59 TyrGluAa 78  
 Db 816 TCTAAATATCTTTTGAAGCTAATCTTCAAAATATGTTAGAGGTCTTTTGTCTAAATCT 875  
 OY 79 ThrIleAaSerGlyThrLeuLeuAlaGlyIleGlyGlyLeuThrSerIleSerGlyPro 98  
 Db 876 GCTGCTATGCTTCAATCTTGTCTACGCTCCAAATACGCTGTGTCATCTATACGACGCTGC 935  
 OY 99 IlegIyIleIlegIyAlaIleIleIleIleIleSerPheGlyThrLeuIleThrValPheTyrPro 118



Db 936 ACAGCTATAGCAGCTATA-----ATACCTATCTTGGCCT 974  
 Qy 119 ALaGluGluGlnAspLeuThrValTrpThrGlnPheIleuMetGlyValIlePheVal 138  
 Db 975 TCT-----CAAGAAATATATCTCCGATTAACCTTAGCTATTTGAGCAACCTT 1028  
 Qy 139 AspThrProLeuThrGluSerIleuGluLeuValIleuGlnThrLeuGluGlyPheArg 158  
 Db 1029 TATTCATTTCAGATCAAGAGGTTAGAGAGATGACATTACCGCATTTGAAAGTTTAA 1088  
 Qy 159 GlnIleuGlnSerTyrAsnThrAlaLeuAspAspTrpArgIleuValArgLeuGln 178  
 Db 1089 GATAGTGAATAATATTGGAATAATGCTTCACATTGGATTAAT----- 1133  
 Qy 179 AlaProGlyLeuProProSerSerAlaLeuGlnGlnAlaIleuThrLeuValArg 198  
 Db 1134 -----AATCTTAATTCAAC-----AACACAACTACTGTGAGCAAAAGA 1172  
 Qy 199 PheGluAsnValHisAsnAspPheIleArgGluIleProGlyPheGluLeuGluThrTyr 218  
 Db 1173 TTCCAAAGGTTAATGTTAGTATTGTAGCGACGATGCTTCTTTAGCAAAAGACTAT 1232  
 Qy 219 LysThrLeuLeuLeuProIleTyrAlaGlnAlaIleAsnPheHisLeuAsnLeuGln 238  
 Db 1233 GAACCAATACGTGTTATCACTTATGACAGCTGACAGTTACATTTCATTACG 1292  
 Qy 239 GlnGlyAlaGluLeuAlaAspGluTrpAsnAlaAspIleHisProSerGlnIleGluPro 258  
 Db 1293 GATGAATTAAGTATGCTGAATAATGCAATTAATGACAAAGGATGACATG----- 1346  
 Qy 259 AsnAlaGlyThrSerAspAspTyrTyrIleuLeuValGluAsnIleProIleTyrSer 278  
 Db 1347 -----CCGGGGATTACTTTACAAAGAAATTAACTAATATGCAATGAATAATATT 1397  
 Qy 279 AsnTyrCysAlaAsnThrTyrArgThrGlyLeuValAsnLeuArgAspGluProAsnMet 298  
 Db 1398 GAGCTTTATTAATGAGTGAATGATAAAGTTTGTCTCTTTTAA-----AGTGTAGTGA 1454  
 Qy 299 LysThrSerIlePheAsnAspTyrArgArgTyrMetThrIleThrValLeuAspThrIle 318  
 Db 1455 AATGTGTAGAAATCAATCAATATGCTGATCTTAACAGCCTGATATGATGCTCAT 1514  
 Qy 319 SerGlnPheSerLeuTyrAspIleValArgTyrArgAspSerIleGlyValIleVal 338  
 Db 1515 TCTCTTTTTCATCATATGACCCAGTTTATTAAGAAAGATTA---AGTGTGAGAT 1571  
 Qy 339 LysGlyIleValAsnGluLeuThrArgGluIleTyrThrGluIleAsnPheAspArg 358  
 Db 1572 -----CTCAGAGAAATTAATATACAGATCCATTAATTAATATCATAGA 1613  
 Qy 359 LeuProGlnLeuArgValGlnProAsnLeuAlaThrMetGluTyrAsnLeuThrArgAla 378  
 Db 1614 GGCAATTCACCTAGAGAGATGAAGCAAAATATACCTTGAGCCAACTTTGTTACTCA 1673  
 Qy 379 SerPheIleValPheSerPheLeuGluGlnPheIlePheTyrThrGluAsnThrAsnPhe 398  
 Db 1674 CTATATACACTTAACCTTATATGTAATATATTTTAAATTCATGCGGACACCAAT--- 1730  
 Qy 399 GlyAsnArgLeuValGlyIleSerAsnArgAspAlaProThrTyrSerAsnThrIleThr 418  
 Db 1731 -----ACTTAATGATTAATTTATCACCG 1751  
 Qy 419 GluThrLeuTyrGlyGluArgThrGlySerProThrThrIleArgProPheGlu 438  
 Db 1752 GACAAATTTTGGACAGGCTTTTGGAAAAACAAGAGAGATATATGAT----- 1802  
 Qy 439 SerTyrIleValSerIleValThrAspArgGlnSerProProValSerProIleGlnPro 458  
 Db 1803 -----AAAGTCCCTGATATCTTAACGATTAATCA----- 1832  
 Qy 459 HisPheIleIleAsnGlnIleGluLeuTyrIleAsn----- 470

Db 1833 -----ATTATTTAATAAATTCGACATATGATATCATTAATGTTATTTAATGTAATG 1886  
 Qy 471 -----GlySerSerAsnAsnThrLeuValTyrSerAlaGlySer 484  
 Db 1887 TACTTTGGTTTGGACGCGGAAAAAGATCAAAATATCAAAATTAATGAGTAGTTCT 1946  
 Qy 485 LeuSerAsnTyrGlnAsnThrThrPhePheGlnPheProArgIleValAspCysAsnLeu 504  
 Db 1947 ACAAGAAATATAC-----ATTAATAATTTGACT--- 1973  
 Qy 505 ValIleAspProGlyCysSerProAsnPheAsnValTyrSerHisIleLeuSerHis--- 523  
 Db 1974 -----CATGCTTACGAGATGTCATGTCATCATAT 2003  
 Qy 524 -----PheSerLeuPheThrTyrSerTyrValIleGlyLeuGlnIleLeuAsp 541  
 Db 2004 CTGATGAAAAAAACAATGTTATCTTT----- 2033  
 Qy 542 ThrGlyValLeuGlyTyrThrHisSerSerValAspArgTyrAsnAlaIleSerAspLys 561  
 Db 2034 -----GCATGACACTTACTATCAATTTCTTGAATAATGAATTAATAATGAT 2081  
 Qy 562 IleIleThrMetIleProAlaIleLeuGlyAsnAsnLeuAspThrAsnSerLysValIle 581  
 Db 2082 ATTATCAACAATCCCTGCTGTAATAAGCTATCAATTAGGTGTCATTCACAACTGAT 2141  
 Qy 582 GluGlyProGlyHisThrGlyValAsnLeuValTyrIleGlnSerGlnGlyArgLeuGlu 601  
 Db 2142 AAAGTCCAGCTCATACTGCTGAGAGACTTAATTAATTAATGAATGATTAATTAAGT 2201  
 Qy 602 IleThrCysGlu-----ThrProAsnSerThrGlnSerTyrPheIleArgLeuArgTyrAla 620  
 Db 2202 ATATCATGTCACACTTATCTTAATGTTACAAAAAATATTTTGTAAGATTCGCTACGCT 2261  
 Qy 621 ThrAsnGlyAlaGlyAsnThrLeuProAsnIleSerLeuThrIleProGlyValIleGly 640  
 Db 2262 ACAATGCTTCTTAATATCAATAGCCCATTAATTAATATTAATCAATTCGA-----GSA 2312  
 Qy 641 IleProGlnArg-----LeuAsnAsnThrPheSerGlyThrAsnTyrAsnAsnLeu 658  
 Db 2313 ATGACACCAACAAGATGATGCTCATGATATCTTTTCTGACACGTTAATAGCAATTTA 2372  
 Qy 659 GlnTyrGlyAspPheGlyTyrPheGlnPheProSerThrValThrLeuProLeuAsnArg 678  
 Db 2373 GAATATCAAAATTTTGGCTAATAAGAAATTTCTAAAGCGTGAACATTAAATCCAAATCA 2432  
 Qy 679 AsnIleProPheIlePheAsnArgAlaAspVal---SerAsnSerIleLeuIleIleAsp 697  
 Db 2433 AGTATATCTCTTAATTAATCGTTCTGATCAAAATTCAAATTCATTTACTCTTGAT 2492  
 Qy 698 LysIleGluPheIleProIleThrSerSerMetHisGlnAsnArgGluValGlnLysLeu 717  
 Db 2493 AGAATTTGAATTTCTACCAATTAATCTCCCTTATACAAAGAAAGTAAAGAACAAATTTA 2552  
 Qy 718 GluThrIleGlnThrIleValIleAsnThrPhePheThrAsnHisThrIleThrLeu 735  
 Db 2553 GAAAAAAATCAGAAAAACGTAAACAAGTTATTTTCAATTAATAATTAAGCGTTG 2606  
 RESULT 15  
 AF132928 4896 bp DNA linear BCT 14-JAN-2000  
 LOCUS Bacillus thuringiensis subsp. finitimus Cry28a1 delta-endotoxin  
 DEFINITION gene, complete cds.  
 ACCESSION AF132928  
 VERSION AF132928.1 GI:4574729  
 KEYWORDS  
 SOURCE  
 ORGANISM  
 . Bacillus thuringiensis serovar finitimus  
 Bacillus thuringiensis serovar finitimus  
 Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus; Bacillus  
 cereus group.  
 REFERENCE 1 (bases 1 to 4896)  
 Wojciechowska, J.A., Lewitin, E., Revina, L.P., Zalunin, I.A. and  
 Chestukhina, G.G.

TITLE Two novel delta-endotoxin gene families cry26 and cry28 from  
JOURNAL *Bacillus thuringiensis* ssp. *finitimus*  
FEB8 Lect. 453 (1-2), 46-48 (1999)  
10403372  
REFERENCE 2 (bases 1 to 4896)  
AUTHORS Wojciechowska, J.A., Lewitin, E.I. and Cheetukhina, G.G.  
TITLE Direct Submission  
JOURNAL Submitted (04-MAR-1999) Laboratory of Protein Chemistry, Institute  
of Microbial Genetics, Dorozhny proezd, Moscow 115545, Russia  
FEATURES  
source  
1. 4896  
/organism="Bacillus thuringiensis serovar finitimus"  
/mol\_type="genomic DNA"  
/strain="VKEM B-1161"  
/sub\_species="finitimus"  
/db\_xref="taxon:29337"  
1129. 4458  
/codon\_start=1  
/translation="11  
/product="Cry28Aa1 delta-endotoxin"  
/protein\_id="AAD24189.1"  
/db\_xref="gi:4574730"  
/translation="MAOTYYKIGVOSTEVSSEIPEPNEVDSPTVAVAGIYVGT  
ILTAFAFVNPGVULISRGTLAPVTPPPEPRPKTKNGQPKHSGEDLNLQTSIRVKE  
IALAHNGFKOVLTTERAFNDKNSPANTARLVSOEPENAFNFSNMPOLOPTY  
DTLLSCYTEALANLNLHQGVQPADQNDQPHSPMLKSGGYDELVLYIEKYIN  
YCTKYHGLNHLKESKIKTWDVNTYREMTLLVLDLVATFPFYDIRPFGVELL  
TREYVTSLDHLTRPGLFTMLSDIELYTESVABGDYLSKESKYTNOPTMKNLY  
GNTNRLSKOILTLTPGEFMTLSINRPOTLAGINKYLSLOKIVFTPKDNEOKY  
FNVNQNPQETTYPNPDYSGNSQKPKHNSHPILTHKLEPAFYSIRALQMTN  
VANSQNLISESVSTQIPLVKAYEVNNSVINGPFGDGLIELRDKSICOKKASLKV  
YASLIFYANNAIVASIDVGSAGVLLQPTFSRKGNNFTIODLVNDYIDKASLKV  
IELPSESEIHILKQEDYERGVLLIDLEKPKIDENYTEMNLEKAKKAVNLYFIN  
ATNALMKQVTDHIDQVANLVESIDLVYAKKIKFTFCISPAKQLSQARNLSPDN  
NNNAENMTANTGVTLIEGDPYKGRALIOASAEDNFPYLVOKIDESILKPYTRY  
OLRGVSGSODLELDVYRGATDYMANPGLDELISYAPINPCEIETRDITTCGAL  
DRKQSNVNSADVDVPQVNGDPHAFPHIDTGTNNRMGLVITKIAITPGQYAT  
FGNLELIELGLSEGLAAQVQKEQKWKNTQKREBAKLYAAKQINQIFPLDSQ  
TKLRPDIFFSNLSADKLVIKIRDYSEVLVIPGLNDLPELENNRIONMIDLYDAR  
NTVTGSEFRNGLANMNASNTREVOIOAHPCMYSLGNMAQVQANLEPDKGYLAVT  
AKKEGINGVYTIIDCANHIDTLTPSSCDSGFTTSSNLELAIVTKILEIFPDTOQIRI  
EIGETRSIFYESVDLIMED"

ORIGIN

Alignment Scores:  
Pred. No.: 4,62e-54 Length: 4896  
Score: 849.00 Matches: 229  
Percent Similarity: 48.84 Conservative: 106  
Best Local Similarity: 33.34 Mismatches: 248  
Query Match: 21.94 Indels: 104  
DB: 1 Gaps: 25

US-10-783-417-2 (1-735) x AF132928 (1-4896)

QY 68 SerAlaapThrThrAlaAlaValSerAlaGlyThrIleValSerGlyThrIleuVala 87  
Db 1210 AGCAAGTGAATACGTCGTGTAAGTAAAGCAGGAGATTGATGCTGGACTACTACAGCA 1269  
QY 88 GlyIleGlyValLeuThrSerIleSerGlyProIleGlyIleIleGlyAlaIleIle 107  
Db 1270 -----GCCCTTGACATTTGTTAATCA-----GGTGGTACTCTTATA 1308  
QY 108 SerPheGlyThrLeuIleThrValPheTrpProAlaGlyGluGlnAsp---LysThrVal 126  
Db 1309 TCATTGGAACTTGCGTCGCCCTTCTTGCGCTGATCCAGAGAGATCCAAAAAATT 1368  
QY 127 TrpThrGlnPheIleLysMetGlyGluIlePheValAspThrProLeuThrGluSerIle 146  
Db 1369 TGGTCACATTTATGAAACACGAGAGACCTTTAAATCAAAACATTTCTACACCTGTA 1428  
QY 147 LysGlnIleuLysLeuGlnThrLeuGluGlyPheArgGlnIleLeuGlnSerTyraThr 166  
Db 1429 AAAAGAAATGACATTGCTCATCTAAATGTTTAAAGATGTAATTAACGTACTATGAAAGA 1488

QY 167 AlaLeuSapAepTrpArgGlyLeuLysArgLeuGlnAlaProGlyLeuProSerSer 186  
Db 1489 GCATTATATGATTGGAAAGACA-----AATCCAAAGTCGA 1521  
QY 187 AlaLeuGlnGlnAlaIleuThrLeuLysIleArgPheGluSapValHisAspPhe 206  
Db 1522 -----AATACTGCGCAGATGGTGTACACAGAGATTTGAAAAGCCCATTTCAATTT 1572  
QY 127 ILeaArgGluIleProGlyPheGluLeuGlnThrTyrrSerThrIleuLeuProIleTy 226  
Db 1573 GTAAGCAATATGCCAACATCCAACTTCCACGATGACACATTTATTAAGTTGCTAT 1632  
QY 227 AlaGlnAlaAlaAspPheIleLeuSapLeuGlnGlnGlyAlaGluLeuAlaAspGlu 246  
Db 1633 ACAGAGCTGCAAAATTTACATTGATTTATTAATCAAGAGTGTACATTCGCCGATCA 1692  
QY 247 TrpAsnAlaAepIleHisProSerGlnIleGluProAsnAlaGlyThrSerAspApyr 266  
Db 1693 TGGAAATGCAATCAACCAATTCACCAATGTTGAAGTCATCAGGTACT-----TAT 1743  
QY 267 TyrIleuLeuLysGluAsnIleProLysTyrrSerAsnTyrrCysAlaAsnThrTyrrArg 286  
Db 1744 TATGACAGCTATTTGGTATATATTGAAGAATATTTATTTATTCACCAAGACATACAT 1803  
QY 287 ThrGlyLeuLysAsnLeuArgAspGluProAsnMetLysTrpSerIlePheAspApyr 306  
Db 1804 AAAGATTTGAATCACTTAAAGAAATCAGAAAAATCACATGGATGCTTATTAACATAT 1863  
QY 307 ArgArgTyrrMetThrIleThrValIleuAspThrIleSerGlnPheSerLeuTyrrAsp 326  
Db 1864 CGTGAGAGAAAGACCTTATATGATTTGATCTTGCGCACTTTCTTTTATGATATTA 1923  
QY 327 LysArgTyrrArgAspSerIleGlyGlyIleGluValLysGlyIleLysAsnGluLeuThr 346  
Db 1924 CGTCGTTT-----CCAAAGAGATGAACTGAATTAACA 1959  
QY 347 ArgGluIleTyrrThrThrGluIleAsnPheAspArgLeuProGlnLeuArgValGlnPro 366  
Db 1960 AGAGAGCTTATACAGATTGATGATCATTTAAACAGACACCA----- 2001  
QY 367 AsnLeuAlaThrMetGluTyrrAsnLeuThrArgAlaSerPheLysLeuPheSerPheLeu 386  
Db 2002 -----GGGCTATTTACTTGCGCTG 2019  
QY 387 GluGlnPheIlePheTyrrThrGluAsnThrAsnPheGlyAsnArgLeuValGlyIleSer 406  
Db 2020 TCAAGATTGATGATTATACACGAGAGTGTGCGACAGCGATTTATTTACAGATAT--- 2076  
QY 407 AsnArgAspAlaProThrTyrrSerAsnThr-----IleThrGluThrLeuTyrrGly 423  
Db 2077 ---CGAGGCTCAATATATATATACGTGAATCAATTTTACAGATGAAAAATATTATGCT 2133  
QY 424 GluArgThrGly---SerProThrThrLysThrIleArgProPheGluSerTyrrLysVal 442  
Db 2134 AATCAAAATAGATTAAAGTACAGCTCATTCATTTATACAGAGGAA---TTTATGACT 2190  
QY 443 SerIleValThrAspArgGlnSerProProValSerProIleGlnProHisPheIleIle 462  
Db 2191 CACTTAAGCATTAACCGTCTTTTCAAAACAAATAGCTGATTAATTAAGTTATACGTTTA 2250  
QY 463 AsnGlnIleGluLeuTyrrLeuAsnGlySerSerAsnAsnThrLeuLysTyrrSerAlaGly 482  
Db 2251 ATTCAAAAAATCGATTTCACAACCTTTAAAAACGATAT-----GATATCAAAAAAAT 2304  
QY 483 GlySerLeuSerAsnTyrr-----GlnAsnThrPhePheGlnPheProArgLys 499  
Db 2305 TTTATGTAGATTAATCAAAATGAACCTTCAGAAAATCTACA----- 2343  
QY 500 LysAspCysAsnLeuValIleAspProGlyCysSerProAsnPheAsnAsnTyrrSerHis 519  
Db 2344 -----AACTATCTTAATGATTATGTGTGCTTCA---AACAGCCAAAAAATTCAAACAT 2391

QY 520 IleuSerHisPheSerIleuPheThrTyrSerTyrValIleGlyLeuGlnLeuGlnIle 539  
DB 2392 AATTATCTCATTTTCATTATCATCCAC-----AAGTTAGAGTTT 2433  
QY 540 LeuAsp-----ThrGlyValLeuGlyTyrThrHisSerSerValAspArgTyr 555  
DB 2434 GGTGAGTATTTTCACCTCTATTTGTCATTAGTTGACACACATATGTGTAACTCCCA 2493  
QY 556 AsnAlaIleSerAspLysIleIleThrMetIleProAlaIleLysGlyAsnAsnLeuAsp 575  
DB 2494 AATTATATATCAGAAAGTGTAGTACACAAATCCCATGTGTAAAGCTTACGAAGTT--- 2550  
QY 576 ThrAsnSerLysValIleGlnGlyProGlyHisThrGlyGlyAsnLeuValTyrLeuGln 595  
DB 2551 ACTAACATTCAGTTATTAAGAGCACAGGTTTACAGTGAGATTATTAAGACTTCGT 2610  
QY 596 SerGlnGlyArgLeuGlnIleThrCysGlnThrProAsnSerThrGlnSerTyrPheIle 615  
DB 2611 GAT-----AAATGTTCTATTAAATGTAAAGCT--AGTCTTTAAATAAATACGCTATA 2661  
QY 616 ArgLeuArgTyrAlaThrAsnGlyAlaGlyAsnThrLeuProAsnIleSerLeuThrIle 635  
DB 2662 AGTCTATTATTATGCTCAATAACGCAATAGCTGTATCAATAGACTAGTGATTCGGA 2721  
QY 636 ProGlyValIleGlyIleProProGlnArgLeuAsnAsnThrPheSerGlyThrAsnTyr 655  
DB 2722 GCAGGAGTTCTA-----TTGCCAACCTACCTTTTCTAGAAAAGGGAAC 2763  
QY 656 AsnAsn-----LeuGlnTyrGlyAspPheGlyTyrPheGlnPheProSer 670  
DB 2764 AATTAATTTTACATTCACAGACCTTAACCTTAAGATTTCATATCATACACTTTTAGTT 2823  
QY 671 ThrValThrLeuProLeuAsnArgAsnIleProPheIlePheAsnArgAlaAsp----- 688  
DB 2824 GATATTGAATTACCGAAGTAGAGAAATTCATATCCATTGAGCGAGAGAGATGATTAT 2883  
QY 689 ValSerAsnSerIleLeuIleIleAspLysIleGlnPheIleProIleThrSerSerMet 708  
DB 2884 GAGGAGGAGGTGATTTTATTATGATTAATTAAGAGTTCAACCTTATAGTGAATAATTAT 2943  
QY 709 HisGlnAsnArgGlnLysGlnLysLeuGlnIleGlnThrLysIleAsnThrPhePhe 728  
DB 2944 -----ACTAATGAATGAATTAGAGAAAGCAAGCAAGCAATGATATATTT 2994  
QY 729 ThrAsnHisThrLysThrLeu 735  
DB 2995 ATTAACGCAACAACGCTTTG 3015

Search completed: January 20, 2006, 20:29:24  
Job time : 9852 secs

**THIS PAGE BLANK (USPTO)**